

AD _____

Award Number: W81XWH-10-1-0488

TITLE: Variations of Human DNA Polymerase Genes as Biomarkers of Prostate Cancer Progression

PRINCIPAL INVESTIGATOR: Nick Makridakis, Ph.D.

CONTRACTING ORGANIZATION: Tulane University
New Orleans, LA 70112

REPORT DATE: July 2011

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE 1 Jul 2011		2. REPORT TYPE Annual		3. DATES COVERED 15 JUN 2010 - 14 JUN 2011	
4. TITLE AND SUBTITLE Variations of Human DNA Polymerase Genes as Biomarkers of Prostate Cancer Progression				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-10-1-0488	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Nick Makridakis, Ph.D. E-Mail: nmakrida@tulane.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Tulane University New Orleans, LA 70112				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Human DNA polymerases beta, eta, and kappa are enzymes that function in repairing damaged DNA. DNA sequencing analysis of 40 prostate tumors identified somatic mutations of these genes in most patients. Biochemical analysis of the polymerase beta missense variations showed that all of these somatic mutations have functional effects. Thus, common polymerase mutations may contribute to prostate cancer progression.					
15. SUBJECT TERMS Biomarker discovery, cancer genetics					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	29	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
Introduction.....	4
<u>Body.....</u>	<u>4-8</u>
Key Research Accomplishments.....	8
Reportable Outcomes.....	8
Conclusion.....	8-9
References.....	9
Appendices.....	10-29

INTRODUCTION

The early molecular events that lead to sporadic prostate cancer progression are largely unknown. Human DNA polymerases beta, eta, and kappa are distributive error-prone enzymes that function in a relatively accurate manner when replicating damaged DNA. We set out to: (1) identify common somatic variants in the human pol beta, pol eta and pol kappa genes in men suffering from prostate cancer. (2) measure the frequency and distribution by tumor stage, tumor grade and patient age of each of the variants identified in (1) in human prostate cancer tissues. (3) To determine the effect of the somatic variants identified in (1) on polymerase function.

BODY

In Task 1, we set out to identify common somatic variants in the human pol beta, pol eta and pol kappa genes in men suffering from prostate cancer. We thus sequenced the coding sequence of the pol β , η and κ genes for somatic mutations in 40 prostate cancer tissues, to test the potential merit of our hypothesis that prostate cancer tissue is commonly mutated in these genes. Briefly, we PCR amplified all exons using DNA isolated from microdissected prostate cancer tissue and then sequenced the PCR products bi-directionally with BigDye chemistry on an ABI sequencer (detailed in Makridakis *et al*, 2009; see *Appendix*). We have completed the pol beta gene, finished all but one exon for pol kappa gene, and have also analyzed 3 exons of the pol eta gene, so far. We identified many somatic mutations in these samples, most of them missense (Table 1, and Makridakis *et al*, 2009).

Although the above strategy is important for the discovery of biomarkers of prostate tumor progression, it would also be highly beneficial to discover common biomarkers of pre-invasive prostate cancer, or even precancerous lesions, for maximum reduction of mortality from this disease. Thus we set out to genotype 40 prostatic intraepithelial neoplasia (PIN; a precursor to prostate cancer) tissues for the presence of the pol beta missense variants we identified (Table 1). PIN and adjacent normal prostate tissues were removed from prostate tissue slides donated by Dr. Zongbing You of Tulane University, by laser-captured microdissection. Genomic DNA was extracted from the PIN tissues by the Picopure[®] DNA Extraction Kit (Molecular Devices; Sunnyvale, CA), and PCR amplified using the same primers that identified the POLB mutations shown on Table 1. However, all PCR amplifications we tested failed, while our positive controls worked (data not shown). We also failed to detect DNA in our PIN tissues using Picogreen (Promega Co; Madison, WI). These data suggest that we did not recover enough DNA from our PIN tissues to get adequate PCR amplification. This is not surprising, since PIN tissue is usually comprised of a single layer of epithelial cells lining the prostate gland.

Given the high number of somatic mutations identified (including many missense mutations; Table 1), we decided to determine the effect of these mutations on biochemical activity (Task 3) prior to establishing their frequency in sporadic prostate tumors (Task 2) (genotyping a high number of mutations makes more sense when one knows that a significant proportion of these mutations are not mere passengers of cancer evolution).

We thus set out to biochemically characterize each missense polymerase beta variant identified in Task 1 (Table 1). We initially focused on missense mutations because they have a higher chance of causing functional effects. Briefly, we reconstructed all missense pol β variations in an appropriate expression vector and then measured the effect of each mutation on enzyme activity, protein expression and fidelity of DNA synthesis *in vitro*, after purification of the respective overexpressed enzymes (detailed in An *et al*, 2011; see *Appendix*).

Gene (accession #)	Patient	DNA change	Type of Change/ location	Mutant peak (%)	Predicted/ known effect
POLB (AF491812.1)	1	g.23912G>A	AG-splice junction	59 %	Deletion of amino acids 184-185
	2	g.25254G>A	E216K	52 %	
	4	g.960C>T	5'-UTR	52 %	
		g.18145T>C	N128N	54%	
	5	g.32481C>T	P261L	100 %	Altered fidelity
		g.32573A>G	T292A	100 %	Altered fidelity
		g.32592T>C	I298T	100 %	
		g.32439G>A*	Intron 12	79 %	
	8	g.11628T>C	Intron 3	27 %	
	10	g.25314A>T	M236L	38 %	Altered fidelity
	11	g.15180G>A	E123K	100%	
	12	g.1444G>C	K27N	60 %	
	13	g.25236C>T	L210L	42 %	
		g.31911C>G**	P242R	40 %	Altered fidelity
	15	g.921C>T	Promoter	100 %	
	20	g.25302G>A	E232K	100 %	
	23	g.12622A>G	AG-splice junction	25 %	Deletion of amino acids 88-90
	24	g.11630T>C	Intron 3	30 %	
	29	g.32521G>T	G274G	35 %	
	30	g.32467T>C	Intron 12	51 %	
		g.32439G>A*	Intron 12	100 %	
POLH (AY388614.1)	1	g.28891G>A	G259R	23 %	
	6	g.29021A>T	Intron 7	46 %	
	20	g.28901G>C	G263A	60 %	XPV
		g.28967C>T	S284F	66 %	
POLK (AY273797.1)	1	c.560T>C	F155S	45%	
	4	c.507T>C	S137S	35%	
		c. 557G>A	G154E	100%	
	6	c.1420C>T	L442F	30%	
	10	c. 1353G>A	E419E	100%	
	12	g.67088C>A	T205K	28 %	
	21	g.66992A>G	Intron 5	68 %	Changes Lariat-A
	27	g.67088C>T	T205I	30 %	
	29	c. 1380G>A	A428A	100%	

		c. 1385A>G	E430G	100%	
	31	c. 2416G>A	G774S	60%	
		c. 2441G>A	C782Y	50%	
	33	c. 2432G>A	C779Y	43%	
		c. 2567G>A	S824N	45%	
		c. 2694T>G	D866E	23%	
	34	c. 425C>T	A110V	20%	
		c. 469G>A	D125N	19%	
		c. 482C>T	A129V	22%	
		c. 2186C>T	S697L	50%	
	35	c. 433G>A	A113T	20%	
		c. 1437G>A	Q447Q	100%	
		c. 1441G>A	E449K	100%	

*Creates a new branch (Lariat-A) site

** Also in constitutional DNA (rs3136797)

Table 1. Somatic mutations detected in polymerase genes in prostate cancer patients.

Notes: Mutations not previously reported are shown in bold. % Mutant indicates the sequencing peak height that corresponds to the mutant (average of forward and reverse sequence). For both mutations that change the invariant splice junction (AG), an in-frame AG exists shortly downstream. Utilization of the alternative AG is predicted to result in the deletions shown in the last column. XPV denotes a pol η residue mutated in an XPV patient. For further details, see text.

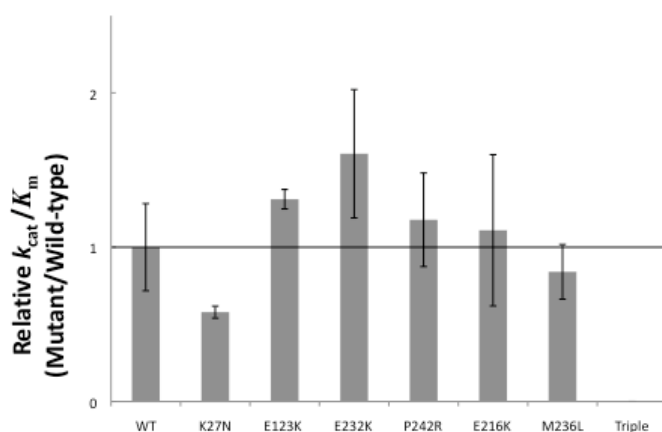


Fig. 1. Influence of POLB variants on catalytic efficiency for dCTP insertion. WT and mutant variants were assayed on a gapped oligo substrate with a templating dG, and the catalytic efficiency (k_{cat}/K_m) were determined from dividing k_{cat} and K_m values obtained by Excel analysis of the inverse plots. These data represent the mean of at least three independent determinations.

In vitro biochemical (polymerase beta) assays for both wild type (WT) and mutant variants were performed on a single-gapped oligonucleotide substrate (the natural polymerase beta substrate) with a templating deoxy-G. The experiments were performed at steady-state conditions for the enzyme, and the catalytic efficiency of each variant was determined, as a measure of enzyme activity. The data comparing WT to mutant efficiencies are shown in Figure 1.

The data presented in Figure 1 indicate that two of the pol β variants significantly reduce catalytic efficiency (K27N and Triple mutant:

P261L/ T292A/ I298T). The remaining variants contain wild type activity.

Some mutations may be active, and yet still affect mutagenicity by altering the fidelity of DNA synthesis. We thus proceeded to characterize the effect of all POLB variations on DNA synthesis fidelity using the same single-gapped oligonucleotide substrate with a templating dG, (detailed in An *et al*, 2011; see *Appendix*). Fidelity measurements were performed with each of the four dNTPs. The data, shown in Figure 2, indicate that the majority of the POLB somatic mutations affect the fidelity of DNA synthesis. Thus, two of the pol β variants assayed reduce catalytic efficiency (Fig. 1), while the remaining five missense mutations alter the fidelity of DNA synthesis (Fig. 2).

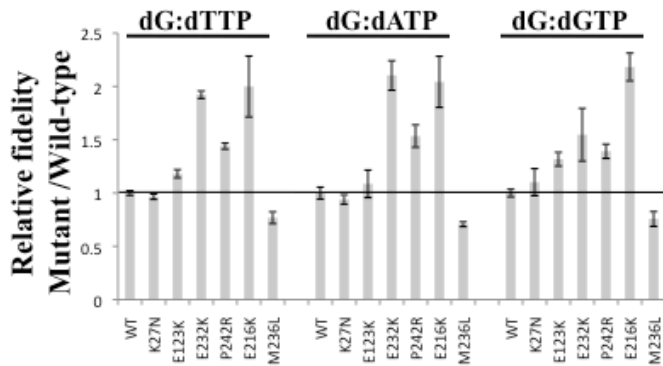


Fig. 2. Relative fidelity of variants enzyme. WT and mutant variants were assayed on single-nucleotide gapped DNA substrate with a templating dG. Fidelity= $[(k_{cat}/K_{m,dNTP})_{correct} + (k_{cat}/K_{m,dNTP})_{incorrect}] / (k_{cat}/K_{m,dNTP})_{incorrect}$. These data represent the mean of at least three independent determinations.

However, fidelity depends on sequence context, and we have thus far analyzed the mutants on a specific sequence context (An *et al*, 2011; see *Appendix*). Therefore, in order to better predict the mutagenic effect of these somatic POLB mutations, we plan to utilize the Forward Mutation Assay (Bebenek *et al*, 2003). This assay relies on scoring errors produced by a DNA polymerase while filling a 407 nucleotide gap on the *lacZ* α gene, and thus can identify all types of

mutations (deletions, insertions, substitutions, etc.) in different sequence contexts. DNA synthesis produces M13mp2 DNA that yields dark blue phage plaques upon introduction into an Escherichia coli α -complementation strain and plating on indicator plates. Errors are scored as light blue or colorless mutant plaques. These plaques are restreaked, to verify that they contain mutations. DNA from independent mutant clones is then isolated and sequenced to define the *lacZ* mutation (Bebenek *et al*, 2003).

Unlike the remaining POLB variants analyzed, the K27N missense mutation is part of the deoxyribose phosphate (dRP) lyase domain of polymerase beta (which is responsible for the removal of the 5'-dRP intermediate formed during base excision repair) (Dalal *et al*, 2008). Thus we compared the dRP lyase activity of K27N to wild type, using the dRP lyase assay (detailed in An *et al*, 2011; see *Appendix*). The results showed that the K27N mutant significantly decreased both the K_m and k_{cat} , resulting in a small decrease in catalytic efficiency, K_{cat}/K_m (Table 2).

Enzyme	K_{cat} (min ⁻¹)	K_m (nM)	K_{cat}/K_m (nM ⁻¹ min ⁻¹)
WT	4.1×10^{-3}	262	1.6×10^{-5}
p.K27N	1.1×10^{-3}	77	1.4×10^{-5}

Table 2. Steady-state kinetic parameters for pol beta deoxyribose lyase activity

Note: the kinetic values presented are calculated as detailed in An *et al*, 2011; see *Appendix*.

KEY RESEARCH ACCOMPLISHMENTS

- Identification of common somatic variants of the human pol β , pol η and pol κ genes in prostate cancer tissue
- Characterization of the biochemical effects of all missense pol β prostate cancer tissue variations identified

REPORTABLE OUTCOMES

Manuscripts

An C, Chen D., and Makridakis NM. (2011). Systematic Biochemical Analysis of Somatic Missense Mutations in DNA Polymerase β Found In Prostate Cancer Reveal Alteration of Enzymatic Function. *Hum Mutat*, 32: 415-423.

Abstracts

Chen D, Mukhopadhyay S., Yadav S., and Makridakis NM. (2011). Polymerase genes, genomic instability and Prostate Cancer, *LCRC Annual Scientific Retreat, Xavier University, New Orleans, LA*.

Presentations

Error-prone polymerase mutations and prostate cancer progression, COBRE/Cancer Genetics group seminar, Tulane University, New Orleans, LA, 08/10

Candidate genes, genomic instability and cancer progression, Genetics seminar, Tulane University, New Orleans, LA, 03/11

Error-prone polymerases, genomic instability and prostate cancer progression, COBRE/ EAB presentation, Tulane University, New Orleans, LA, 03/11

Funding applied for based on work supported by this award

Polymerase genes and genomic instability in prostate cancer progression, NCI/R01, 10/2010

CONCLUSION

Overall, our DNA sequencing analysis so far shows that 73% of the 40 prostate cancer patients we examined had somatic substitutions in at least one of the polymerase genes tested (Table 1). Therefore, error-prone polymerase genes are commonly mutated in prostate tumor tissue.

Furthermore, biochemical analysis of the pol beta missense variations showed that all of these somatic mutations have functional effects, either by reducing enzymatic activity (Fig. 1), or by altering the fidelity of DNA synthesis (Fig. 2). Thus, this research has uncovered the novel concept that functional alteration of error-prone DNA polymerases may be common and important for prostate cancer progression.

"So what?"

Given the prevalence of functionally important mutations of polymerase beta in prostate tumors, we propose that screening for these somatic mutations in prostate biopsies may be an important biomarker for the detection of prostate cancer. Characterization of these mutations in additional tumors is expected to determine the potential benefit of those mutations as biomarkers of disease outcome. Similar analysis should be performed for the other two polymerase genes, as proposed in our plan. This approach may eventually lead to the characterization of a novel mechanism for prostate carcinogenesis and progression, involving the accumulation of mutations resulting from a defective polymerase genotype.

REFERENCES

An C, Chen D., and Makridakis NM. (2011). Systematic Biochemical Analysis of Somatic Missense Mutations in DNA Polymerase β Found In Prostate Cancer Reveal Alteration of Enzymatic Function. *Hum Mutat*, 32: 415-423.

Bebenek K, Garcia-Diaz M, Blanco L, Kunkel TA. (2003). The frameshift infidelity of human DNA polymerase lambda. Implications for function. *J Biol Chem*. 278: 34685-90.

Dalal S, Chikova A, Jaeger J, Sweasy JB. (2008). The Leu22Pro tumor-associated variant of DNA polymerase beta is dRP lyase deficient. *Nucleic Acids Res*. 36:411-22.

Makridakis NM, Ferraz L, and Reichardt J KV. (2009). Genomic Analysis of Cancer Tissue Reveals that Somatic Mutations Commonly Occur in a Specific Motif, *Hum Mutat*, 30: 39-48.

APPENDICES:

List:

1. Makridakis NM, Ferraz L, and Reichardt J KV. (2009). Genomic Analysis of Cancer Tissue Reveals that Somatic Mutations Commonly Occur in a Specific Motif, *Hum Mutat*, 30: 39-48.
2. An C, Chen D., and Makridakis NM. (2011). Systematic Biochemical Analysis of Somatic Missense Mutations in DNA Polymerase β Found In Prostate Cancer Reveal Alteration of Enzymatic Function. *Hum Mutat*, 32: 415-423.

Genomic Analysis of Cancer Tissue Reveals That Somatic Mutations Commonly Occur in a Specific Motif

Nick M. Makridakis,^{1*} Lúcio Fábio Caldas Ferraz,¹ and Juergen K.V. Reichardt²

¹Department of Biochemistry and Molecular Biology, Institute for Genetic Medicine, Keck School of Medicine, University of Southern California (USC), Los Angeles, California

²Plunkett Chair of Molecular Biology (Medicine), University of Sydney, Camperdown, Australia

Communicated by David E. Goldgar

Received 12 November 2007; accepted revised manuscript 1 April 2008.

Published online 11 July 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/humu.20810

ABSTRACT: Somatic mutations are hallmarks of cancer progression. We sequenced 26 matched human prostate tumor and constitutional DNA samples for somatic alterations in the SRD5A2, HPRT, and HSD3B2 genes, and identified 71 nucleotide substitutions. Of these substitutions, 79% (56/71) occur within a WKVnRRRnVWK sequence (a novel motif we call THEMIS [from the ancient Greek goddess of prophecy]: W = A/T, K = G/T, V = G/A/C, R = purine (A/G), and n = any nucleotide), with one mismatch allowed. Literature searches identified this motif with one mismatch allowed in 66% (37/56) of the somatic prostate cancer mutations and in 74% (90/122) of the somatic breast cancer mutations found in all human genes analyzed. We also found the THEMIS motif with one allowed mismatch in 88% (23/26) of the ras1 gene somatic mutations formed in the sensitive to skin carcinogenesis (SENCAR) mouse model, after induction of error-prone DNA repair following mutagenic treatment. The high prevalence of the motif in each of the above mentioned cases cannot be explained by chance ($P < 0.046$). We further identified 27 somatic mutations in the error-prone DNA polymerase genes pol η , pol κ , and pol β in these prostate cancer patients. The data suggest that most somatic nucleotide substitutions in human cancer may occur in sites that conform to the THEMIS motif. These mutations may be caused by “mutator” mutations in error-prone DNA polymerase genes.

Hum Mutat 30, 39–48, 2009.

© 2008 Wiley-Liss, Inc.

KEY WORDS: cancer; somatic; mutation; consensus; error-prone DNA polymerase; SRD5A2; HPRT; HSD3B2

Introduction

Cancer is thought to evolve through the accumulation of somatic mutations in specific genes, depending on tumor type [Vogelstein and Kinzler, 2004]. These mutations are caused by a combination of environmental and heritable factors [Lichtenstein et al., 2000]. To date scientists have been unable to identify a common motif at the sites of these somatic mutations, suggesting that these somatic events have distinct molecular etiologies, depending, among other factors, on the individual and the type of tumor. We tested the hypothesis that there is a common motif at the sites of somatic mutations in prostate cancer tissue, a tumor type with poorly understood etiology.

We screened the HSD3B2, SRD5A2, and HPRT genes for somatic mutations. The SRD5A2 gene (MIM# 607306) encodes the steroid 5 α -reductase type II enzyme that reduces testosterone (T) to dihydrotestosterone (DHT), the most active androgen in the prostate [Cheng et al., 1993]. Thus activating 5 α -reductase mutations may contribute to prostate tumor development. A possible contribution of the SRD5A2 gene to prostate tumor progression was also proposed based on our findings of de novo somatic events at the SRD5A2 locus [Akalu et al., 1999a]. More recently, we reported the biochemical characterization of somatic SRD5A2 mutations in human prostate cancer tissue [Makridakis et al., 2004], including mutations that increase enzyme activity. The HSD3B2 gene (MIM# 109715) also presents an attractive androgen metabolic candidate gene for prostate cancer risk and progression. The type II 3 β -hydroxysteroid dehydrogenase enzyme encoded by the HSD3B2 gene is mainly expressed in androgenic tissues [Labrie et al., 1992; Lachance et al., 1991], and initiates DHT inactivation [Cheng et al., 1993]. However, if a specific mechanism generates somatic mutations at distinct motifs, then it should do so even in genes expected to be unrelated to tumor progression, such as the human HPRT gene. The ubiquitously expressed HPRT gene (MIM#s 308000 and 300322) is located on the X chromosome [Pai et al., 1980] and encodes hypoxanthine-guanine phosphoribosyl transferase, an enzyme essential for the purine salvage pathway, responsible for 90% of nucleic acid biosynthesis in normal cells [Zoref and Sperling, 1979].

The Supplementary Material referred to in this article can be accessed at <http://www.interscience.wiley.com/jpages/1059-7794/suppmat>.

Current affiliation for Lúcio Fábio Caldas Ferraz: Centro de Biologia Molecular e Engenharia Genética, University of Campinas, São Paulo, Brazil.

Current affiliation for Nick M. Makridakis: Tulane Cancer Center, Tulane University, New Orleans, Louisiana.

*Correspondence to: N.M. Makridakis, Tulane Cancer Center, Tulane University, 1430 Tulane Avenue, SL-68, New Orleans, LA 70112.

E-mail: nmakrida@tulane.edu

Contract grant sponsor: National Cancer Institute (NCI); Grant number: P01 CA108964 (Project 1); Contract grant sponsor: American Cancer Society–Institutional Research Grant (ACS-IRG); Grant number: 58-007-45.

We report here 71 somatic mutations in these three genes. Analysis of the nucleotide sequence surrounding these mutations revealed that these alterations often occur within a novel motif we call THEMIS (from the ancient Greek goddess of prophecy). In search of a potential mutagenic mechanism we found that the THEMIS motif commonly occurs at the sites of somatic mutations induced by error-prone (EP) DNA repair in the sensitive to skin carcinogenesis (SENCAR) mouse skin model following mutagenic treatment [Chakravarti et al., 2000, 2001].

EP DNA repair (also called translesion synthesis [Goodman, 2002]) involves DNA polymerases such as pol η (MIM# 603968) and pol κ (MIM# 605650), which are much more accurate when replicating through specific types of DNA damage than undamaged DNA [Pages and Fuchs, 2002]. EP DNA polymerases have been proposed to play a role in cancer etiology [see Kunkel, 2003]. To date, the only known example of an EP DNA polymerase causing human cancer comes from pol η [Kunkel, 2003]; constitutional DNA mutations that inactivate pol η are associated with a high rate of skin cancers in patients suffering from XPV (xeroderma pigmentosum variant [Johnson et al., 1999]). In the absence of pol η , ultraviolet (UV) radiation-induced pyrimidine dimers are bypassed in a manner that generates the mutations which lead to skin cancer [Kunkel, 2003], perhaps by another EP DNA polymerase [Pages and Fuchs, 2002]. Thus EP polymerase mutations can result in multiple tumor-inducing mutations (mutator phenotype [Loeb et al., 2003]) given the right type (and amount) of environmental exposure.

Human pol β (MIM# 174760) is a DNA polymerase essential for base excision repair (BER) [Sancar et al., 2004]. BER is one of the major pathways of DNA repair that removes oxidized and alkylated bases from DNA [Friedberg, 2003]. Pol β is not a classic error prone polymerase, yet it causes 67 times more substitution errors than mammalian pol δ [Kunkel, 2003]. Pol β is also involved in meiotic recombination [Goodman, 2002] and repair of double-stranded DNA breaks through the process of non-homologous end-joining [Wilson and Lieber, 1999]. Interestingly, both the pol κ and pol β genes are located in chromosomal regions known to be lost during prostate cancer progression (5q and 8p, respectively; e.g., Visakorpi et al. [1995]). Thus, somatic loss of these polymerases may contribute to prostate tumor progression.

We screened the pol β , η , and κ genes for somatic prostate cancer mutations, to test the hypothesis that EP polymerases are commonly mutated in prostate cancer tissue. We report 27 somatic mutations in these EP polymerase genes in the same prostate cancer tissues that have additional somatic mutations in the other analyzed genes reported here (i.e., HSD3B2, HPRT, and SRD5A2).

Materials and Methods

Tumor Specimens

We analyzed 26 patients of Caucasian background with prostatic adenocarcinoma (for further description see Akalu et al. [1999a] and Makridakis et al. [2004]). These patients underwent radical prostatectomy at the USC Norris Comprehensive Cancer Center. Prostate tissue and blood were collected from each patient. Tumors were staged according to the tumor, nodes, metastases (TNM) staging system (see Supplementary Table S1a; available online at <http://www.interscience.wiley.com/jpages/1059-7794/suppmat>) [Schroder et al., 1992]. Local Institutional Review Board (IRB) approval was obtained before study initiation.

Microdissection and DNA Extraction

Specimens were formalin fixed, embedded in paraffin, sectioned, and transferred on microscopic slides, where they were deparaffinized and stained with hematoxylin and eosin. Selected populations of carcinoma cells were microdissected and tumor DNA was then extracted from the microdissected cells using a method reported by us earlier [Akalu and Reichardt, 1999b]. As control, normal (constitutional) DNA was extracted either from microdissected normal cells adjacent to the tumor or from peripheral blood leukocytes (or both).

Molecular Analysis

PCR

The entire coding region of the HSD3B2 gene, together with the exon-intron splicing junction boundaries, the putative promoter region, and the 5' and 3' untranslated regions (UTRs), were amplified by PCR reactions using sets of primers as previously described [Rheume et al., 1992; Simard et al., 1993]. Purified (desalted) oligonucleotides were obtained from IDT (Coralville, IA). Reactions were performed with the polymerases AmpliTaq Gold (Applied Biosystems [ABI], Foster City, CA) or HotStart Taq (Qiagen, Valencia, CA) with their corresponding PCR buffer. Reactions carried out with HotStart Taq had an additional reagent (5 \times Q-solution) provided by the manufacturer. The reaction mixture consisted of 10 to 20 ng of DNA, 1 \times PCR buffer, 1.5 mM MgCl₂, 200 μ M dNTPs, 0.1 to 0.2 μ M of each forward and reverse primer, 1 \times Q-solution (when necessary), and 2.5 units of polymerase and sterile water, in a final volume of 50 μ L. The reaction was then covered with mineral oil and subjected to thermal cycling in a RoboCycler Gradient 40 (Stratagene, La Jolla, CA) under the following conditions: an initial pre-PCR heat step of 95°C for 2 min (AmpliTaq Gold) or 15 min (HotStart Taq), 50 cycles of denaturation at 95°C for 1.5 min, annealing at 58 to 70°C for 1.5 min, and elongation at 72°C for 1.5 min. This was followed by a final extension step at 72°C for 10 min. PCR products were purified with the QIAquick Gel Extraction Kit (Qiagen) and then sequenced.

Exons 7–8 of the HPRT gene were PCR amplified as a single 379-bp fragment as described above, except that we used primers described elsewhere [Liu et al., 2003]. Sequencing analysis of the HPRT gene was performed as in the following paragraphs. The original analysis of the SRD5A2 gene is described in Makridakis et al. [2004].

The EP polymerase genes pol β , pol η , and pol κ were PCR-amplified as described above, except that we used the primers shown in Supplementary Table S1b.

Sequencing

Sequencing reactions consisted of 20 to 60 ng of the purified PCR product, 3.2 pmol of each PCR primer, 4 μ L of 5 \times sequencing reaction buffer (ABI), 4 μ L of ABI PRISMTM Dye Terminator Cycle Sequencing Ready Reaction Mix (ABI), and sterile water in a total volume of 20 μ L. Reactions were submitted to the following thermal cycle: 96°C for 30 s, 50°C for 15 s, and 60°C for 4 min, for a total of 30 cycles. The PCR reactions were then purified according to the manufacturer's recommendations and submitted to electrophoresis. Nucleotide sequences were collected on either an ABI PRISM 377 or a 3100 Automated DNA Sequencer (ABI). Results were processed with the ABI PRISM Sequence Navigator software (ABI). Nucleotide substitutions were

identified and quantitated by ABI PRISM Factura Feature Identification software with “identify heterozygote limit” set at 10% (ABI). Each nucleotide substitution was confirmed by at least three independent PCR sequencing analyses.

Nucleotide changes are numbered according to the appropriate genomic reference sequence, as noted, according to journal guidelines (www.hgvs.org/mutnomen).

Microsatellite analysis

The complex dinucleotide repeat in intron 3 of the HSD3B2 gene was amplified by PCR using a fluorescent primer. Tumor DNA and its matched normal DNA were amplified by PCR using the set of primers previously described [Verreault et al., 1994]. In some cases, the alternative reverse primer 5'-TGGACCTATGTTTGTGTGG-3' was utilized, yielding fragments 96 bp shorter than those described by Verreault et al. [1994]. The forward primer [Verreault et al., 1994] was labeled on the 5' end with the fluorescent dye TETTM. PCR reactions were performed with HotStart Taq as described above and submitted to the following procedure: initial denaturation step for 15 min at 95°C, 50 cycles with a denaturation step at 95°C for 1 min, annealing at 62°C for 1 min, and extension at 72°C for 1 min, followed by a final extension step at 72°C for 10 min. PCR products were loaded on an ABI PRISM 377 Automated DNA Sequencer together with an internal size standard (Genescan-500TM tetramethylrhodamine [TAMRA]; ABI, Foster City, CA), according to the manufacturer's recommendations. Genotyping analyses were performed using the Genescan software (ABI). Experiments were performed in triplicate.

Statistics

All P values were calculated using the chi-square test, with the Yates correction when appropriate (for one degree of freedom).

Determination of the Actual Target of the THEMIS Motif

For each gene, we analyzed the sequence that contained the somatic mutations (usually the sequence that was PCR-amplified, except for androgen receptor (AR) and TP53, for which we used the cDNA sequence with the addition of short patches of intronic sequence, for the mutations that were on intron-exon boundaries). The analysis was performed using a software program that finds known motifs in DNA sequence imported by the user, called “dna-pattern (strings)” (available at: http://rsat.ulb.ac.be/rsat/dna-pattern_form.cgi). We ran this program online using both 0 and 1 allowed substitutions from the motif, and analyzed the results.

Results

Our prostate cancer samples were previously described [Akalu et al., 1999a]. Supplementary Table S1a summarizes patient information including age and tumor stage as well as all results concerning microsatellite instability (MSI) and nucleotide substitutions in the HSD3B2 gene. These results are presented in detail in the sections that follow.

Somatic Mutations in the HSD3B2 and HPRT Genes in Prostate Cancer Tissue

If there is a common somatic mutation motif in prostate cancer tissue, then it should be ubiquitous. We thus decided to screen both genes containing potential “driver” mutations (such as HSD3B2; Makridakis et al. [2005]) and passenger mutations (such as HPRT). Initially, the tumor DNA of each patient was screened

for somatic alterations, by automated DNA sequencing. Once nucleotide alterations were detected, the constitutional DNA of the same patient was also analyzed to examine if the event was somatic. Somatic events were quite common. In fact, among the 26 patients studied, only five had no detectable nucleotide HSD3B2 alterations in their tumor (Supplementary Table S1a). We detected 38 single-nucleotide somatic mutations total in this gene: one deletion and 37 nucleotide substitutions (Supplementary Table S1a). The HPRT gene was also investigated for somatic alterations in prostate cancer by DNA sequencing: 17 somatic nucleotide substitutions were identified (Table 1). Although only a fraction of the HPRT gene was screened, 9 out of the 26 patients analyzed harbor somatic nucleotide substitutions (Table 1).

Nucleotide Sequence Context and Nature of the Somatic Mutations

The high number of somatic mutations identified in prostate cancer tissue allowed us to determine the sequence context of these nucleotide substitutions. This analysis revealed that most mutations occur within a WKVnRRRnVWK (THEMIS) motif when one mismatch is allowed ($W = A/T$, $K = G/T$, $V = G/A/C$, $R = \text{purine (A/G)}$; $n = \text{any nucleotide}$; total number of $n = 0-2$ nucleotides; the underline indicates the position of the mutated base; Table 2).

A total of 30 out of the 38 (79%) nucleotide alterations detected in the HSD3B2 gene fit the THEMIS motif with up to one mismatch (Supplementary Table S2). The actual target of the motif (the RRR sequences occurring within the WKVnRRRnVWK context with one allowed mismatch) covers 60% of the HSD3B2 sequence analyzed (see Materials and Methods for calculation). Thus, 22.8 (60%) of the HSD3B2 mutations are expected to occur in the motif (with up to one mismatch), compared to the 30 (79%) observed ($P = 0.026$). We then tested the occurrence of the motif in the HPRT and SRD5A2 somatic mutation sites: 15 out of 17 (88%) of the somatic HPRT mutations and 11 out of 16 (69%) of the previously reported [Makridakis et al., 2004] somatic SRD5A2 mutations fit the motif with one allowed mismatch (Table 1; Supplementary Table S3). The actual motif target (with one allowed mismatch) covers 68% of the HPRT and 57% of the SRD5A2 sequences analyzed (see Materials and Methods). Thus 20.6 (out of the 33) HPRT/ SRD5A2 mutations are expected to fit the motif with one allowed mismatch, compared to the 26 (15+11) observed ($P = 0.05$). Comprehensive analysis of all the HPRT, HSD3B2, and SRD5A2 somatic mutations indicates that the number of mutations that fit the motif with up to one mismatch (56) is significantly higher ($P = 0.005$) than expected (43.4), based on the number of motifs that exist in these genes.

Of the 54 somatic SRD5A2 and HSD3B2 substitutions, 94% are transitions (Supplementary Table S1a) [Makridakis et al., 2004]. In contrast, only 52% of the 63 constitutional DNA variations found in these genes in controls and in patients with HSD3B2 deficiency [Makridakis et al., 2004; Pang, 2001] (and data not shown) are transitions ($P = 0.000003$). Moreover, 88% of the 17 HPRT somatic substitutions are transitions (Table 1). In contrast, only 51% of the 59 listed HPRT SNPs (<http://snpper.chip.org/bio/export-sequence/20797>) are transitions ($P = 0.019$). Thus transitions are significantly more common at the somatic mutation sites.

The inclusion of potential driver mutations (such as the missense SRD5A2 mutations) [Makridakis et al., 2004] in our analysis may result in a motif that is biased toward variants that survive the tumor selection process rather than the underlying mutation process. To examine this possibility we analyzed the presumed neutral

Table 1. Somatic Nucleotide Substitutions in the HPRT Gene*

Patient	DNA change	Type of change/location ^a	Forward sequence	Reverse sequence	Mutant nucleotide peak (%)				
					Sequence context ^b				
1	g.39835C>T ^c	P169S	15	25	gTG		GGG	TC	CTT
	g.39897C>T	Intronic	22	15	AGA	T	GGt	TA	AAT
	g.39985T>C	Intronic	34	25	AGA	TT	tAA		AAG
4	g.40051C>T ^c	P184S	19	34	TGt	CT	GGA		ATT
	g.39835C>T ^c	P169S	19	29	gTG		GGG	TC	CTT
	g.40051C>T ^c	P184S	23	34	TGt	CT	GGA		ATT
6	g.39835C>T ^c	P169S	14	24	gTG		GGG	TC	CTT
	g.40051C>T ^c	P184S	21	32	TGt	CT	GGA		ATT
	g.40055A>G	D185G	28	35	TTC	C	AGA	C	AAG
11	g.40073A>G	Y191C	28	20	gGA	T	AtG	CC	CTT
	g.40091A>G	E197G	29	22	ATG		AAt	A	CTT
	g.39959A>G	Intronic	32	38	TTG	T	AGA		GAG
16	g.39890A>G ^c	Intronic	33	32	TTA	A	AtG		ATT
	g.40086T>C	Y195Y	44	41	TGA	CT	AtA		ATG
	g.39929G>A	Intronic	18	19	TGA	AA	tGG		CTT
20	g.40000C>T ^c	Intronic	38	42	ATA		AAG		AAa
	g.39859G>A	D177N	81	69	AGC	C	AGA		CTG
	g.39939G>T	Intronic	50	66	ATA	A	tGG		CTT
21	g.40011T>C	Intronic	37	30	ATC	TA	AAt		GAT
	g.39907C>T ^c	Intronic	65	89	TTA		GGt	TA	AAG
	g.39971G>T ^c	Intronic	23	33	TGG		cAA		ATG

*The HPRT nucleotide changes are numbered according to the genomic reference sequence (GenBank: M26434.1).

^aIn the "Type of change" column, missense mutations are indicated using the single letter code.

^bIn the "Sequence context" column, the nucleotide sequence is displayed using the DNA strand that rendered the best fit to the THEMIS motif. The mutated base is indicated in bold. Lowercase letters indicate mismatches within the consensus motif. The spacer nucleotides are underlined. Sequence contexts that conform to the motif are shaded. The mutant peak intensity is reported as percentage of total peak intensity.

^cMutations that fit the motif only in the nontranscribed strand (the rest of the mutations fit the motif in the transcribed strand).

Table 2. Most Somatic Mutations in Cancer fit the THEMIS Motif*

Breast		Prostate	
Gene	% Fit	Gene	% Fit
CHEK2	90	H-RAS, K-RAS, N-RAS	50
CSNK1E	82	TP53	67
GATA3	67	PTEN	100
LMO4	100	AR	68
CDH-1	84	HSD3B2	79
DLG1	67	SRD5A2	69
UBR-5	80	HPRT	88
BRCA2	100		
PHB	60		
PTEN	64		
TP53	67		

*The percentage of somatic mutations that fit the WKVnRRRnVWK motif (with one mismatch allowed) is given, by gene and type of cancer. Bold indicates data reported in this report. The rest of the data was analyzed from the published literature. For details and references see Results.

substitutions (i.e., those that do not change the protein structure) separately: this data suggests that neutral substitutions in all three genes fit the motif well ($P = 0.005$; Supplementary Table S4a) and that in fact there is no significant difference in the motif-fitting ratios among presumed drivers and presumed neutral substitutions ($P = 0.3$; Supplementary Table S4b).

Nucleotide substitution rates vary according to sequence context and clearly depend on the nearest neighbor nucleotides (e.g., Lunter and Hein [2004]). We examined whether an established model that predicts nucleotide substitution rates based on sequence context (the dinucleotide substitution model; Lunter and Hein [2004]) fits the observed mutation spectra better

than the THEMIS motif. The data, presented in Supplementary Table S5, suggests that the dinucleotide substitution model predicts a much different distribution than the observed ($P < 0.0001$). Thus the observed somatic mutations in all the genes we examined do not fit the expected distribution based on dinucleotide substitution model, but instead frequently fit the THEMIS motif.

The THEMIS Motif in the Cancer Literature

The discovery of the THEMIS motif prompted us to examine the published literature for other kinds of mutations that may fit this motif. This search revealed that the motif is found (with one allowed mismatch) in 88% (23/26) of the *ras1* gene mutations detected in the SENCAR mouse skin cancer model after benzo[α]pyrene [Chakravarti et al., 2000] or estradiol-3, 4-quinone [Chakravarti et al., 2001] treatment (Supplementary Table S6; and data not shown). Benzo[α]pyrene is a known etiologic agent of both skin and lung cancer [Chakravarti et al., 2000]. These mouse skin H-ras mutations were the result of EP DNA repair [Chakravarti et al., 2000, 2001]. Of the 26 *ras1* gene mutations, 24 were also transitions (Supplementary Table S6; data not shown). The actual motif target covers 62% of the *ras1* sequence analyzed (see Materials and Methods). Thus, 16.2 (62%) of the *ras1* mutations are expected to occur in the motif (with one mismatch), compared to the 23 (88%) observed ($P = 0.01$). The frequency of the *ras1* mutations that fit the exact motif (without mismatches) is also higher than expected (35% vs. 21%) (Supplementary Fig. S1), but this trend is not significant.

Additional literature searches identified the THEMIS motif in 66% (37/56) of the somatic prostate cancer mutations that we found in TP53, H-ras, K-ras, N-ras, PTEN, and the AR gene (Table 2). Of

these mutations, 68% are transitions. Examination of the somatic mutations in the AR gene database [Gottlieb et al., 2004] revealed that 68% of the mutations fit this motif (with up to one mismatch) and 80% are transitions. Moreover, one-half of the most common somatic mutations (in prostate cancer) that activate the H-ras, K-ras, and N-ras oncogenes, at codons 12, 13, and 61, also fit the THEMIS motif with one mismatch allowed. Regarding somatic substitutions in the TP53 gene in prostate cancer [Chi et al., 1994], there is a prevalence of transitions over transversions and 67% (14/21) of the mutations fit this motif with up to one mismatch. The most common somatic PTEN mutation, found in 1 out of 6 prostate cancer tissues [Pesche et al., 1998], also fits this motif. To test the significance of these observations, we calculated the expected number of the AR gene mutations that fit the THEMIS motif based on the actual motif target (see Materials and Methods) in the AR gene, the most commonly mutated gene in this dataset: 42 out of the 62 (68%) somatic AR mutations [Gottlieb et al., 2004] fit the motif with one mismatch, compared to 32.8 (53%) expected ($P = 0.027$). The frequency of the somatic AR mutations that fit the THEMIS motif exactly (without mismatches) is also higher than expected (16 observed vs. 10.4 expected), but this trend does not reach statistical significance ($P = 0.08$). Moreover, 80% of the somatic prostate cancer mutations in the AR gene [Gottlieb et al., 2004] are transitions, while only 57% of the constitutional AR SNPs (<http://snpper.chip.org/bio/export-sequence/20479>) are transitions ($P = 0.0054$).

After detecting the THEMIS motif in the most common human malignancy in U.S. males [Jemal et al., 2007], we decided to also examine the most common malignancy in females, breast cancer [Jemal et al., 2007]. The data presented in Supplementary Table S7 shows that 74% (90/122) of all somatic breast cancer mutations found in the literature, fit this motif with up to one mismatch. The list of analyzed genes includes CHEK2 [Staalesen et al., 2004; Sullivan et al., 2002], CSNK1E [Fuja et al., 2004], GATA3 [Usary et al., 2004], LMO4 [Sutherland et al., 2003], CDH-1 [Becker et al., 1999; Berx et al., 1996], DLG1 [Fuja et al., 2004], EDD/hHYD (UBR-5) [Fuja et al., 2004], BRCA2 [Lancaster et al., 1996], PTEN [Kurose et al., 2002], TP53 [Sullivan et al., 2002; Kurose et al., 2002; Thorlacius et al., 1993], and PHB [Sato et al., 1992, 1993] (Table 2). To test the significance of this finding, we calculated the expected number of somatic mutations that fit the THEMIS motif in the most commonly mutated gene in this dataset (TP53), based on the actual motif target in TP53: 29 out of the 43 (67%) somatic TP53 mutations fit the motif with one mismatch allowed, compared to 23.5 expected (see Materials and Methods). This higher than expected incidence does not reach statistical significance ($P = 0.089$). However, 10 out of the 43 (23%) somatic TP53 mutations fit the motif exactly, compared to 5.6 (13%) expected, and this higher incidence is significant ($P = 0.046$). Moreover, 58% of all somatic breast cancer mutations reported in these genes are transitions (Supplementary Table S7). Of the TP53 SNPs (<http://snpper.chip.org/bio/export-sequence/7966>), 55% are transitions. The difference in the transition frequency between breast cancer and constitutional TP53 substitutions is not significant.

Recently, a GA pattern (or TC in the opposite DNA strand; the mutated base is underlined) was identified at the sites of somatic mutations in the protein kinase gene family in breast cancer [Stephens et al., 2005]. Interestingly this pattern emerged mostly from two breast cancer samples thought to display mutator phenotype [Stephens et al., 2005]. Since this pattern is similar to the purine core of the WKVNRRRnVWK motif, we decided to analyze these mutations for the presence of the THEMIS motif: we found that 59 out of the 88 (67%) breast cancer mutations fit the

THEMIS motif with one allowed mismatch. In contrast, only 8 out of the 71 prostate cancer mutations that we report here occur in the GA motif (8.9/71; i.e., 12.5% expected; $P = 0.75$). The GA motif is overrepresented though in the breast cancer mutation dataset (Supplementary Table S7): 25 out of the 122 mutations fit the GA motif, compared to 15.3 expected ($P = 0.01$). Thus both breast and prostate cancer mutations occur in the context of a purine-rich core motif, but in breast cancer this core is often GA.

Somatic Mutations in the pol β , pol η , and pol κ Genes in Prostate Cancer Tissue

The significant presence of the THEMIS motif at the sites of somatic ras1 mutations induced by EP DNA repair suggests that such EP DNA polymerases may be involved in the etiology of (at least some) of the somatic mutations that fit this motif. Accordingly, we decided to sequence selected exons from each of the EP DNA polymerase genes pol β , pol η , and pol κ for somatic mutations in our prostate cancer tissues, to test the hypothesis that prostate cancer tissue bears common somatic mutations in these genes. This preliminary analysis identified somatic mutations in all three genes, but the gene with most (and more prevalent) mutations was pol β . Therefore, we screened the complete coding sequence of the pol β gene in these patients. The result of these analyses in all polymerase genes is shown in Table 3: We identified 27 somatic mutations in these 26 samples, 14 of which were missense substitutions, nine were silent or intronic substitutions, two substitutions changed splice acceptor (AG) sites, one was in the promoter region, and one was in the 5'-UTR. The G31438A intronic substitution is recurrent in Patients 5 and 30 (Table 3). The P242R missense substitution was also present in the constitutional DNA of Patient 13, but with altered prevalence (40% average mutant peak in the tumor tissue [Table 3] compared to 60% average mutant peak in the constitutional DNA [data not shown]). Thus an additional somatic event may have occurred in the tumor in this pol β site. Overall, among 26 patients analyzed, 19 patients (73%) have somatic mutations in an EP polymerase gene and 16 patients (61%) have somatic mutations in pol β .

Somatic Instability in Prostate Cancer Tissue

Since the HSD3B2 gene was most commonly "hit" in our sequencing analysis of these prostate cancer tissues, we decided to search for other kinds of genomic instability at this locus. Accordingly, we investigated 20 informative patients for MSI and loss of heterozygosity (LOH) in the complex dinucleotide repeat in intron 3 of the HSD3B2 gene [Devgan et al., 1997], by comparing tumor and matched normal DNA. The results show that 70% of the patients analyzed have LOH/ MSI in the HSD3B2 locus (Supplementary Table S1a). The most common findings were MSI, manifested as follows: five cases with contraction of the alleles of the tumor DNA, four cases with expansion, and three cases with a combination of both contraction and expansion (data not shown). Three patients had LOH in the tumor tissue (Supplementary Table S1a).

Discussion

Somatic Mutations Are Commonly Found in Prostate Cancer Tissue

We report a high number of somatic prostate cancer mutations in all three "target" genes studied. A total of 38 nucleotide

Table 3. Somatic Mutations Detected in Polymerase Genes in Prostate Cancer Patients*

Gene (accession #)	Patient	DNA change	Type of change/location	Mutant peak (%)	Predicted/known effect
POLB (AF491812.1)	1	g.23912G>A	AG-splice junction	59	Deletion of amino acids 184–185
	2	g.25254G>A	E216K	52	
	4	g.960C>T	5'-UTR	52	
		g.18145T>C	N128N	54	
	5	g.32481C>T	P261L	100	Altered fidelity
		g.32573A>G	T292A	100	Altered fidelity
		g.32592T>C	I298T	100	
		g.32439G>A ^a	Intron 12	79	
	8	g.11628T>C	Intron 3	27	
	10	g.25314A>T	M236L	38	Altered fidelity
	11	g.15180G>A	E123K	100	
	12	g.1444G>C	K27N	60	
	13	g.25236C>T	L210L	42	
		g.31911C>G ^b	P242R	40	Altered fidelity
	15	g.921C>T	Promoter	100	
	20	g.25302G>A	E232K	100	
	23	g.12622A>G	AG-splice junction	25	Deletion of amino acids 88–90
	24	g.11630T>C	Intron 3	30	
	29	g.32521G>T	G274G	35	
	30	g.32467T>C	Intron 12	51	
		g.32439G>A ^a	Intron 12	100	
POLH (AY388614.1)	1	g.28891G>A	G259R	23	XPV
	6	g.29021A>T	Intron 7	46	
	20	g.28901G>C	G263A	60	
		g.28967C>T	S284F	66	
POLK (AY273797.1)	12	g.67088C>A	T205K	28	Changes Lariat-A
	21	g.66992A>G	Intron 5	68	
	27	g.67088C>T	T205I	30	

*For both mutations that change the invariant splice junction (AG), an in-frame AG exists shortly downstream. Utilization of the alternative AG is predicted to result in the deletions shown in the last column. XPV denotes a pol η residue mutated in an XPV patient. For further details, see Results.

^aCreates a new branch (Lariat-A) site.

^bAlso in constitutional DNA (rs3136797).

alterations were detected in the HSD3B2 gene in 80% of the 26 patients. We also identified 16 somatic SRD5A2 substitutions in 60% of the patients [Makridakis et al., 2004] and 17 somatic HPRT substitutions in 35% of the patients (Table 1). Collectively, we found a total of 71 somatic mutations in these three genes in prostate cancer. The high rate of somatic events identified, even in the HPRT gene, suggests that there may be generalized genomic instability, at least in some of these tumors. As summarized in Table 4, four patients harbor somatic substitutions in all three “target” genes examined, and most of these substitutions fit the THEMIS motif. Significantly, all of these patients have additional somatic mutations in an EP polymerase gene (Table 4).

Nucleotide Sequence Context of the Somatic Substitutions

Analysis of the nucleotide sequence that surrounds each HSD3B2, SRD5A2, and HPRT gene mutation showed that 79% of the 71 somatic mutation sites fit the THEMIS motif, with up to one mismatch ($P = 0.005$; Table 2). The prevalence of this motif at somatic mutation sites led us to hypothesize that it is common, perhaps even universal. Review of the published literature identified this motif (with up to one mismatch) in 66% of all somatic prostate cancer mutations and 74% of all somatic breast cancer mutations ($P = 0.027$ for prostate; $P = 0.046$ for breast cancer). Moreover, 67% of the somatic breast cancer mutations thought to result through a mutator phenotype in the protein kinase gene family [Stephens et al., 2005] fit the THEMIS motif with one allowed mismatch. Our analyses, therefore, suggest that

most somatic mutations in the most common human malignancies fit the THEMIS motif with up to one mismatch.

Analysis of the type of mutations found in the HSD3B2, SRD5A2, and HPRT genes showed that transitions are significantly more common than expected ($P < 0.019$). This finding was confirmed among the other somatic prostate cancer mutations that we found in the published literature. However, analysis of the breast cancer mutations obtained from the literature shows that their transition frequency is not significantly higher compared to constitutional substitutions. Thus, although both prostate and breast cancer mutations fit the THEMIS motif, only in prostate cancer are there more transitions than expected by chance. Moreover, the GA motif is overrepresented only among breast cancer mutation sites, not prostate cancer. These data suggest that similar yet distinct molecular etiologies exist between the generation of somatic mutations in the prostate and breast. For example, both prostate and breast cancer mutations may be caused by EP repair, but the polymerase or the carcinogen involved may be different.

The THEMIS motif includes two “spacer segments” that are of variable length, from 0 to 2 nucleotides. The spacer segments are often in the 1-1 permutation (i.e., one nucleotide gap in each side of the mutation) but the data reported here is only statistically significant when the spacer is variable (0–2). In addition, the degenerate nature of the motif inevitably has the effect that the same mutation site often fits more than one permutation (e.g., both 0-1 and 1-1). This fact makes it difficult to calculate exact probabilities for each specific permutation. Variable spacers have

Table 4. Patients With Somatic Nucleotide Substitutions in All “Target” Genes Analyzed (HSD3B2, SRD5A2, and HPRT Genes)[†]

Patient	Gene ^a	DNA change	Sequence context ^a			
1	HPRT	g.39835C>T ^c	gTG	GGG	TC	CCT
		g.39897C>T	AGA	T	GGt	TA
		g.39985T>C	AGA	TT	tAA	AAG
		g.40051C>T ^c	TGt	CT	GGA	ATT
	SRD5A2	g.888G>A	TGC	C	AGc	CcG
		g.1890G>A	AcA		AGG	TGG
	HSD3B2	g.8774C>T ^c	TGG	GT	GGA	GTT
	POLB	g.23912G>A	ATC	A	cAG	GTG
	POLH	g.28891G>A	GTC	TT	GGA	G
	6	HPRT	g.39835C>T ^c	gTG	GGG	TC
g.40051C>T ^c			TGt	CT	GGA	ATT
SRD5A2		g.888G>A	TGC	C	AGc	CcG
		g.1294T>C ^c	AGA		AGG	CAG
HSD3B2		g.1551A>G	AGC		AGG	AgG
		g.1571G>A	TcA		GAG	GAT
		g.1622T>C	TcC		AAG	GCC
		g.1671A>G	AGt		AAA	CTT
POLH		g.29021A>T	TTT	TT	AAA	ATC
11		HPRT	g.40055A>G	TTC	C	AGA
	g.40073A>G		gGA	T	AtG	CC
	g.40091A>G		ATG		AAt	A
	g.2019T>C		CGC		AGc	CC
	SRD5A2	g.2019T>C	CGC		AGc	CC
		g.8089G>A	AGA		AGG	
	HSD3B2	g.8174G>A	TGG		GGA	AG
		g.15180G>A	cTA		GAA	G
	POLB	g.15180G>A	AGC	C	AGA	CTG
	21	HPRT	g.39859G>A	AGC		AGA
g.39939T>G			AGC		AAt	TAT
g.40011T>C			ATC	TA	AAt	
g.888G>A			TGC	C	AGc	
SRD5A2		g.888G>A	TGC	C	AGc	
		g.1914G>A	cTG		GAG	CC
HSD3B2		g.1927C>T ^c	AcC		GAG	GA
		g.8006A>G	AGG		AAA	T
POLK		g.8577T>C ^c	AGG	T	GAA	
		g.66992A>G	ATT	T	AAA	

[†]Refer to Table 1 for legend. Sequence contexts that conform to the THEMIS motif are shaded. GenBank sequences: HPRT, M26434.1; SRD5A2, L03843.1; HSD3B2, M77144.1; POLB, AF491812.1; POLH, AY388614.1; and POLK, AY273797.1.

^aPolymerase mutations are in bold.

been reported before at motif sites in nature. For example, the spacer between the -35 and -10 consensus sequences of the Sigma A protein binding site varies between 16 and 18 nucleotides [Helmann, 1995], while the spacer between the two boxes of the Sigma B binding site is 13 to 15 nucleotides long [Petersohn et al., 1999]. The biological mechanism that allows this spacer variability is unknown, but we speculate that the wide and flexible active site [Perlow-Poehnelt et al., 2004] of the EP polymerases may be responsible: it may accommodate the invariable 9 nucleotides of the THEMIS motif in various manners, e.g., with 0-1, 1-1, or 1-0 spacers. Y-family polymerases (such as pol η and pol κ) actually have two partially overlapping active sites [Chandani and Loehler, 2007], a finding that may also contribute to the spacer variability: the two active sites may have different binding preferences.

EP DNA Polymerases: A Role in the Origin of Somatic Mutations?

The existence of a frequently mutated somatic motif suggests that many mutations in the most common forms of human cancer may have similar molecular etiology. The THEMIS motif is

overrepresented among somatic mutation sites induced by EP repair in the SENCAR mouse skin model following mutagenic treatment (Supplementary Table S5 and data not shown; $P = 0.01$). These data parallel our findings in prostate/breast cancer tissue. Thus, we propose that many somatic mutations in many types of human cancer (such as breast, prostate, skin, and lung) are caused by EP DNA repair following carcinogenic exposure.

To test this model we decided to first screen the EP DNA polymerase genes POLB (pol β), POLH (pol η), and POLK (pol κ) for somatic mutations that may have in turn caused the high number of motif mutations in our prostate cancer samples. We report a total of 27 somatic mutations in the three EP polymerase genes in our 26 samples (Table 3; only the pol β gene was fully sequenced). Four somatic pol β mutations have been previously reported in prostate cancer, but that study was conducted with fewer samples (12 cases) of Asian background [Dobashi et al., 1994]. Most (52%) of the somatic mutations reported here are missense, while two change AG splice junctions and another mutation changes a lariat-adenine to guanine (Table 3). The two splice-junction mutations are predicted to result in deletions of two to three amino acids, at the minimum. Most (59%) of the 27 somatic mutations are prevalent in the tumor, suggesting that they may be “drivers.”

X-ray crystallography studies of human pol β may shed a light on the potential role of the mutated residues: proline-261 forms a hydrogen bond with glutamine-264 [Pelletier et al., 1996], while threonine-292 and methionine-236 are both involved in template binding [Sawaya et al., 1997; Bose-Basu et al., 2004]. Disruption of the hydrogen bond between residues 261 and 264 has been proposed to result in the mutator phenotype displayed by the previously reported prostate cancer-associated I260M pol β mutation [Dalal et al., 2005]. Thus the P261L mutation may also disrupt this hydrogen bond and result in prostate tumorigenesis. Moreover, both T292A and M236L substitutions are predicted to destroy the hydrogen bond between the respective residues and the DNA template, affecting DNA synthesis fidelity. Methionine-236 and proline-242 are in the “flexible loop,” a part of pol β that functions to position the primer and that has been shown to contain several residues that cause a mutator phenotype when mutated [Dalal et al., 2004]. The P242R mutant resulted in similar activity and a four-fold lower mutation rate than the normal pol β when assayed in vitro by a herpes simplex virus thymidine kinase forward mutation assay [Hamid and Eckert, 2005]. Thus, the decreased prevalence of the P242R mutant in the tumor compared to the adjacent normal tissue of patient 13 (see Results) may translate into higher mutagenicity in the tumor. Therefore, several of the somatic pol β mutations that we report here may cause a mutator phenotype [Loeb et al., 2003]. Future functional studies ought to examine this prediction.

One of the missense pol η mutations is in glycine-263 and both missense pol κ mutations are in threonine-205 (Table 3). The homologous residue of pol κ threonine-205 is pol η threonine-122 [Boudsocq et al., 2002]. Missense mutations of both pol η residues 122 and 263 were reported in XPV patients [Broughton et al., 2002]. Thus three of the somatic mutations that we identified are in (or are in residues homologous to) pol η residues previously associated with XPV. This finding suggests that these mutations may play a role in carcinogenesis. Yet if some XPV patients are born with a mutation of the same pol η codon that is associated with prostate cancer etiology, then why do these patients get only skin and not prostate cancer? A potential explanation is different environmental exposure: XPV patients are

Table 5. Models for the Potential Etiology of the Observed Somatic Mutation Motifs*

DNA sequence motif/type of mutation		Potential mutagen involved in DNA damage	Potential polymerase involved in DNA damage repair and/or mutation generation
WKVnRRRnVWK	RR on the strand that fits the motif	PAH [Meehan et al., 1977] Cisplatin [Masutani et al., 2000] AAF-dG 8-oxo-dG	pol β [Venugopal et al., 2005]/pol η [Ogi et al., 2002] ^a pol η [Masutani et al., 2000] pol η [Masutani et al., 2000] pol η [Haracska et al., 2000]
	YY on the other strand	UV radiation [Akiyama et al., 1996] Alkylating agent [Prakash and Sherman, 1973]	pol η [Yu et al., 2001] EP polymerase (e.g., pol β [Sobol and Wilson, 2001])
Transition			

*These models involve the action of a specific environmental mutagen acting in conjunction with an error-prone polymerase to cause specific kinds of mutations (or at specific sites). Pol β is not a "classic" error prone polymerase, yet it causes 67 times more substitution errors than mammalian pol δ [Kunkel, 2003]. The models presented in this table are not mutually exclusive (e.g., a transition may also occur in the WKVnRRRnVWK motif). Bold letters indicate the potential sites of the altered nucleotide. For further details see Discussion.

^aThe choice of the polymerase involved may depend on the type of the PAH adduct that is present (depurinating or stable).

PAH, polycyclic aromatic hydrocarbons (e.g., benzo[α]pyrene); AAF-dG, acetylaminofluorene-deoxyguanosine; 8-oxo-dG, 8-oxo-deoxyguanosine; EP, error-prone; UV, ultraviolet.

inevitably exposed to sunlight, but not necessarily to prostate cancer-inducing mutagens (the prostate is not exposed to sunlight and pol η may be important for repairing other types of damage relevant to the prostate; Table 5). Alternatively, the tumor type specificity may result from the exact nature of the mutation.

Multiple factors determine the mutagenic potential of DNA damage; among them, the choice of the DNA repair machinery evoked to repair the lesions [Pages and Fuchs, 2002] and the nucleotide sequence context in which a lesion occurs [Beard et al., 2002; Wei et al., 1995]. We identified a motif around the sites of human prostate and breast cancer mutations and mouse skin cancer mutations induced by EP DNA repair. A model for the generation of these mutations (Table 5) may involve the action of environmental mutagens acting in conjunction with (mutant) EP polymerase genotypes to cause specific kinds of mutations and/or at specific sites. A mutant DNA polymerase may result in mutations directly, through decreased fidelity, or indirectly, through the use of a "nonoptimal" polymerase for each DNA lesion (as in the XPV paradigm; Pages and Fuchs, 2002). The presence of a motif around the mutation sites can result from either the mutagenic tendencies of specific polymerases or the binding requirements of specific mutagens. The type of EP polymerase(s) involved in the etiology of these mutations may be inferred from the known specificities of these enzymes (Table 5), if the type of carcinogenic exposure is known for each patient.

Other mechanisms leading to the generation of the motif mutations are also possible, such as aberrant mismatch repair or BER. In fact, one of the human polymerases that we propose to be involved in this process, pol β (Table 5), is essential for short-patch BER [Sancar et al., 2004]. Pol β -deficient mouse fibroblasts do not show the high amounts of A:T to G:C transitions seen in pol β -positive cells following benzo[α]pyrene treatment [Venugopal et al., 2005], suggesting that pol β is responsible for most of these transitions. Moreover, the fact that the THEMIS motif is degenerate may indicate the presence of more than one mechanism (e.g., more than one polymerase or type of damage) with distinct sequence requirements. Alignment of more than one specific sequence motif may result in a degenerate motif.

In summary, we report here a DNA sequence motif commonly found in prostate cancer mutation sites, termed THEMIS (WKVnRRRnVWK). We extended this finding to human breast and mouse skin cancer and suggest that the THEMIS motif may be the result of EP DNA polymerase-induced mutations in these tumors.

Acknowledgments

We thank Dr. Myron Goodman (USC) and Dr. Christine Hackel (São Paulo, Brazil) for their helpful comments, and Dr. Joe Hacia, Dr. Nunzio Bottini, Troy Phipps, and Dr. Hooman Allayee (USC) for critically reading this manuscript.

References

- Akalu A, Elmajian DE, Highshaw RA, Nichols PW, Reichardt JKV. 1999a. Somatic mutations at the SRD5A2 locus encoding prostatic steroid 5 α -reductase during prostate cancer progression. *J Urol* 161:1355–1358.
- Akalu A, Reichardt JKV. 1999b. A reliable PCR amplification method for microdissected tumor cells obtained from paraffin-embedded tissue. *Genet Anal* 15:229–233.
- Akiyama N, Alexander D, Aoki Y, Noda M. 1996. Characterization of mutations induced by 300 and 320 nm UV radiation in a rat fibroblast cell line. *Mutat Res* 372:119–131.
- Beard WA, Shock DD, Yang XP, DeLauder SF, Wilson SH. 2002. Loss of DNA polymerase beta stacking interactions with templating purines, but not pyrimidines, alters catalytic efficiency and fidelity. *J Biol Chem* 277:8235–8242.
- Becker KF, Reich U, Schott C, Becker I, Bex G, van Roy F, Hofler H. 1999. Identification of eleven novel tumor-associated E-cadherin mutations. *Mutation in Brief* #215. Online. *Hum Mutat* 13:171.
- Berx G, Cleton-Jansen AM, Strumane K, de Leeuw WJ, Nollet F, van Roy F, Cornelisse C. 1996. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. *Oncogene* 13:1919–1925.
- Bose-Basu B, DeRose EF, Kirby TW, Mueller GA, Beard WA, Wilson SH, London RE. 2004. Dynamic characterization of a DNA repair enzyme: NMR studies of [methyl- ^{13}C]methionine-labeled DNA polymerase beta. *Biochemistry* 43:8911–8922.
- Boudsocq F, Ling H, Yang W, Woodgate R. 2002. Structure-based interpretation of missense mutations in Y-family DNA polymerases and their implications for polymerase function and lesion bypass. *DNA Repair* 1:343–358.
- Broughton BC, Cordonnier A, Kleijer WJ, Jaspers NG, Fawcett H, Raams A, Garritsen VH, Stary A, Avril MF, Boudsocq F, Masutani C, Hanaoka F, Fuchs RP, Sarasin A, Lehmann AR. 2002. Molecular analysis of mutations in DNA polymerase eta in xeroderma pigmentosum-variant patients. *Proc Natl Acad Sci USA* 99:815–820.
- Chakravarti D, Mailander PC, Cavalieri EL, Rogan EG. 2000. Evidence that error-prone DNA repair converts dibenzo[a,h]pyrene-induced depurinating lesions into mutations: formation, clonal proliferation and regression of initiated cells carrying H-ras oncogene mutations in early preneoplasia. *Mutat Res* 456:17–32.
- Chakravarti D, Mailander PC, Li KM, Higginbotham S, Zhang HL, Gross ML, Meza JL, Cavalieri EL, Rogan EG. 2001. Evidence that a burst of DNA depurination in SENCAR mouse skin induces error-prone repair and forms mutations in the H-ras gene. *Oncogene* 20:7945–7953.
- Chandani S, Loehrer EL. 2007. Molecular modeling benzo[a]pyrene N2-dG adducts in the two overlapping active sites of the Y-family DNA polymerase Dpo4. *J Mol Graph Model* 25:658–670.

- Cheng E, Lee C, Aaronson SA, Grayhack J. 1993. Endocrinology of the prostate. In: Lepor H, Lawson RE, editors. Prostate diseases. Philadelphia: W.B. Saunders. p 57–71.
- Chi SG, deVere White RW, Meyers FJ, Siders DB, Lee F, Gumerlock PH. 1994. p53 in prostate cancer: frequent expressed transition mutations. *J Natl Cancer Inst* 86:926–933.
- Dalal S, Kosa JL, Sweasy JB. 2004. The D246V mutant of DNA polymerase beta misincorporates nucleotides: evidence for a role for the flexible loop in DNA positioning within the active site. *J Biol Chem* 279:577–584.
- Dalal S, Hile S, Eckert KA, Sun KW, Starcevic D, Sweasy JB. 2005. Prostate-cancer-associated I260M variant of DNA polymerase beta is a sequence-specific mutator. *Biochemistry* 44:15664–15673.
- Devgan SA, Henderson BE, Yu MC, Shi C-Y, Pike MC, Ross RK, Reichardt JKV. 1997. Genetic variation of 3 β -hydroxysteroid dehydrogenase type II in three racial/ethnic groups: implications for prostate cancer risk. *Prostate* 33:9–12.
- Dobashi Y, Shuin T, Tsuruga H, Uemura H, Torigoe S, Kubota Y. 1994. DNA polymerase beta gene mutation in human prostate cancer. *Cancer Res* 54:2827–2829.
- Friedberg EC. 2003. DNA damage and repair. *Nature* 421:436–440.
- Fuja TJ, Lin F, Osann KE, Bryant PJ. 2004. Somatic mutations and altered expression of the candidate tumor suppressors CSNK1 epsilon, DLG1, and EDD/hHYD in mammary ductal carcinoma. *Cancer Res* 64:942–951.
- Goodman ME. 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* 71:17–50. Review.
- Gottlieb B, Beitel LK, Wu JH, Trifiro M. 2004. The androgen receptor gene mutations database (ARDB): 2004 update. *Hum Mutat* 23:527–533.
- Hamid S, Eckert KA. 2005. Chemotherapeutic nucleoside analogues and DNA polymerase β . *Proc Am Assoc Cancer Res* 46:1369.
- Haracka L, Yu SL, Johnson RE, Prakash L, Prakash S. 2000. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase ϵ . *Nat Genet* 25:458–461.
- Helmann JD. 1995. Compilation and analysis of *Bacillus subtilis* sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res* 23:2351–2360.
- Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. 2007. Cancer statistics. *CA Cancer J Clin* 57:43–66.
- Johnson RE, Kondratik CM, Prakash S, Prakash L. 1999. hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science* 285:263–265.
- Kunkel TA. 2003. Considering the cancer consequences of altered DNA polymerase function [Review]. *Cancer Cell* 3:105–110.
- Kurose K, Gilley K, Matsumoto S, Watson PH, Zhou XP, Eng C. 2002. Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. *Nat Genet* 32:355–357.
- Labrie F, Simard J, Luu-The V, Belanger A, Pelletier G. 1992. Structure, function and tissue-specific gene expression of 3 β -hydroxysteroid dehydrogenase/5-ene-4-ene isomerase enzymes in classical and peripheral intracrine steroidogenic tissues. *J Steroid Biochem Mol Biol* 43:805–826.
- Lachance Y, Luu-The V, Verreault H, Dumont M, Rheame E, Leblanc G, Labrie F. 1991. Structure of the human type II 3 β -hydroxysteroid dehydrogenase/ δ^5 - δ^4 isomerase (3 β -HSD) gene: adrenal and gonadal specificity. *DNA Cell Biol* 10:701–711.
- Lancaster JM, Wooster R, Mangion J, Phelan CM, Cochran C, Gumbs C, Seal S, Barfoot R, Collins N, Bignell G, Patel S, Hamoudi R, Larsson C, Wiseman RW, Berchuck A, Iglehart JD, Marks JR, Ashworth A, Stratton MR, Futreal PA. 1996. BRCA2 mutations in primary breast and ovarian cancers. *Nat Genet* 13:238–240.
- Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K. 2000. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343:78–85.
- Liu SX, Cao J, An H. 2003. Molecular analysis of Tripterygium hypoglaucum (level) Hutch-induced mutations at the HPRT locus in human promyelocytic leukemia cells by multiplex polymerase chain reaction. *Mutagenesis* 18:77–80.
- Loeb LA, Loeb KR, Anderson JP. 2003. Multiple mutations and cancer. *Proc Natl Acad Sci USA* 100:776–781.
- Lunter G, Hein J. 2004. A nucleotide substitution model with nearest-neighbour interactions. *Bioinformatics* 20(Suppl 1):i216–i223.
- Makridakis NM, Akalu A, Reichardt JKV. 2004. Identification and characterization of somatic steroid 5 α -reductase (SRD5A2) mutations in human prostate cancer tissue. *Oncogene* 23:7399–7405.
- Makridakis NM, Buchanan K, Tilley W, Reichardt JK. 2005. Androgen metabolic genes in prostate cancer predisposition and progression. *Front Biosci* 10:2892–2903.
- Masutani C, Kusumoto R, Iwai S, Hanaoka F. 2000. Mechanisms of accurate translesion synthesis by human DNA polymerase ϵ . *EMBO J* 19:3100–3109.
- Meehan T, Straub K, Calvin M. 1977. Benzo[alpha]pyrene diol epoxide covalently binds to deoxyguanosine and deoxyadenosine in DNA. *Nature* 269:725–727.
- Ogi T, Shinkai Y, Tanaka K, Ohmori H. 2002. Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. *Proc Natl Acad Sci USA* 99:15548–15553.
- Pages V, Fuchs RP. 2002. How DNA lesions are turned into mutations within cells? [Review]. *Oncogene* 21:8957–8966.
- Pai GS, Sprenkle JA, Do TT, Mareni CE, Migeon BR. 1980. Localization of loci for hypoxanthine phosphate dehydrogenase and glucose-6-phosphate dehydrogenase and biochemical evidence of nonrandom X chromosome expression from studies of a human X-autosome translocation. *Proc Natl Acad Sci* 77:2810–2813.
- Pang S. 2001. Congenital adrenal hyperplasia owing to 3 β -hydroxysteroid dehydrogenase deficiency. *Endocrinol Metab Clin North Am* 30:81–99.
- Pelletier H, Sawaya MR, Wolffe W, Wilson SH, Kraut J. 1996. Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* 35:12742–12761.
- Perlow-Poehnell RA, Likhterov I, Scicchitano DA, Geacintov NE, Broyde S. 2004. The spacious active site of a Y-family DNA polymerase facilitates promiscuous nucleotide incorporation opposite a bulky carcinogen-DNA adduct: elucidating the structure-function relationship through experimental and computational approaches. *J Biol Chem* 279:36951–36961.
- Pesche S, Latil A, Muzeau F, Cussenot O, Fournier G, Longy M, Eng C, Lidereau R. 1998. PTEN/MMAC1/TEP1 involvement in primary prostate cancers. *Oncogene* 16:2879–2883.
- Petersohn A, Bernhardt J, Gerth U, Höper D, Koburger T, Völker U, Hecker M. 1999. Identification of sigma(B)-dependent genes in *Bacillus subtilis* using a promoter consensus-directed search and oligonucleotide hybridization. *J Bacteriol* 181:5718–5724.
- Prakash L, Sherman F. 1973. Mutagenic specificity: reversion of iso-1-cytochrome c mutants of yeast. *J Mol Biol* 79:65–82.
- Rheame E, Simard J, Morel Y, Mebarki F, Zachmann M, Forest MG, New MI, Labrie F. 1992. Congenital adrenal hyperplasia due to point mutations in the type II 3 β -hydroxysteroid dehydrogenase gene. *Nat Genet* 1:239–245.
- Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K, Linn S. 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* 73:39–85.
- Sato T, Saito H, Swensen J, Olifant A, Wood C, Danner D, Sakamoto T, Takita K, Kasumi F, Miki Y. 1992. The human prohibitin gene located on chromosome 17q21 is mutated in sporadic breast cancer. *Cancer Res* 52:1643–1646.
- Sato T, Sakamoto T, Takita K, Saito H, Okui K, Nakamura Y. 1993. The human prohibitin (PHB) gene family and its somatic mutations in human tumors. *Genomics* 17:762–764.
- Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. 1997. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 36:11205–11215.
- Schroder FH, Hermanek P, Denis L, Fair DR, Gospodarowicz MK, Pavone-Maculoso M. 1992. The TNM classification of prostate cancer. *Prostate* 4:129–138.
- Simard J, Rhéaume E, Sanchez R, Laflamme N, de Launoit Y, Luu-The V, van Seters AP, Gordon RD, Bettendorf M, Heinrich U, Moshang T, New M, Labrie F. 1993. Molecular basis of congenital adrenal hyperplasia due to 3 β -hydroxysteroid dehydrogenase deficiency. *Mol Endocrinol* 7:716–728.
- Sobol RW, Wilson SH. 2001. Mammalian DNA beta-polymerase in base excision repair of alkylation damage. *Prog Nucleic Acid Res Mol Biol* 68:57–74.
- Staalesen V, Falck J, Geisler S, Bartkova J, Borresen-Dale AL, Lukas J, Lillehaug JR, Barteck J, Lønning PE. 2004. Alternative splicing and mutation status of CHEK2 in stage III breast cancer. *Oncogene* 23:8535–8544.
- Stephens P, Edkins S, Davies H, Greenman C, Cox C, Hunter C, Bignell G, Teague J, Smith R, Stevens C, O'Meara S, Parker A, Tarpey P, Avis T, Barthorpe A, Brackenbury L, Buck G, Butler A, Clements J, Cole J, Dicks E, Edwards K, Forbes S, Gorton M, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jones D, Kosmidou V, Laman R, Lugg R, Menzies A, Perry J, Petty R, Raine K, Shepherd R, Small A, Solomon H, Stephens Y, Tofts C, Varian J, Webb A, West S, Widaa S, Yates A, Brasseur F, Cooper CS, Flanagan AM, Green A, Knowles M, Leung SY, Looijenga LH, Malkowicz B, Pierotti MA, Teh B, Yuen ST, Nicholson AG, Lakhani S, Easton DF, Weber BL, Stratton MR, Futreal PA, Wooster R. 2005. A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. *Nat Genet* 37:590–592.
- Sullivan A, Yuille M, Repellin C, Reddy A, Reelfs O, Bell A, Dunne B, Gusterson BA, Osin P, Farrell PJ, Yulug I, Evans A, Ozelik T, Gasco M, Crook T. 2002. Concomitant inactivation of p53 and Chk2 in breast cancer. *Oncogene* 21:1316–1324.
- Sutherland KD, Visvader JE, Choong DY, Sum EY, Lindeman GJ, Campbell IG. 2003. Mutational analysis of the LMO4 gene, encoding a BRCA1-interacting protein, in breast carcinomas. *Int J Cancer* 107:155–158.
- Thorlacius S, Borresen AL, Eyfjord JE. 1993. Somatic p53 mutations in human breast carcinomas in an Icelandic population: a prognostic factor. *Cancer Res* 53:1637–1641.
- Usary J, Llaca V, Karaca G, Presswala S, Karaca M, He X, Langerød A, Kåresen R, Oh DS, Dressler LG, Lønning PE, Strausberg RL, Chanock S, Borresen-Dale AL,

- Perou CM. 2004. Mutation of GATA3 in human breast tumors. *Oncogene* 23:7669–7678.
- Venugopal V, Mailander PC, Chakravarti D. 2005. The role of DNA polymerase β in inducing cytotoxicity and mutations in dibenzo[a,l]pyrene-treated mouse fibroblasts. *Proc Am Assoc Cancer Res* 46:3059.
- Verreault H, Dufort I, Simard J, Labrie F, Luu-The V. 1994. Dinucleotide repeat polymorphisms in the HSD3B2 gene. *Hum Mol Genet* 3:384.
- Visakorpi T, Hyytinen E, Koivisto P, Tanner M, Keinänen R, Palmberg C, Palotie A, Tammela T, Isola J, Kallioniemi OP. 1995. In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* 9:401–406.
- Vogelstein B, Kinzler KW. 2004. Cancer genes and the pathways they control. *Nat Med* 10:789–799.
- Wei D, Maher VM, McCormick JJ. 1995. Site-specific rates of excision repair of benzo[a]pyrene diol epoxide adducts in the hypoxanthine phosphoribosyl-transferase gene of human fibroblasts: correlation with mutation spectra. *Proc Natl Acad Sci USA* 92:2204–2208.
- Wilson TE, Lieber MR. 1999. Efficient processing of DNA ends during yeast nonhomologous end joining. Evidence for a DNA polymerase beta (Pol4)-dependent pathway. *J Biol Chem* 274:23599–23609.
- Yu SL, Johnson RE, Prakash S, Prakash L. 2001. Requirement of DNA polymerase eta for error-free bypass of UV-induced CC and TC photoproducts. *Mol Cell Biol* 21:185–188.
- Zoref E, Sperling O. 1979. Increased de novo purine synthesis in cultured skin fibroblasts from heterozygotes for the Lesch-Nyhan syndrome: a sensitive marker for carrier detection. *Hum Hered* 29:64–68.

Systematic Biochemical Analysis of Somatic Missense Mutations in DNA Polymerase β Found in Prostate Cancer Reveal Alteration of Enzymatic Function

Chang Long An, Desheng Chen, and Nick M. Makridakis*

Department of Epidemiology and Tulane Cancer Center, Tulane University, New Orleans, Louisiana

Communicated by Dominique Stoppa-Lyonnet

Received 28 April 2010; accepted revised manuscript 3 January 2011.

Published online 1 March 2011 in Wiley Online Library (www.wiley.com/humanmutation). DOI 10.1002/humu.21465

ABSTRACT: DNA polymerase β is essential for short-patch base excision repair. We have previously identified 20 somatic pol β mutations in prostate tumors, many of them missense. In the current article we describe the effect of all of these somatic missense pol β mutations (p.K27N, p.E123K, p.E232K, p.P242R, p.E216K, p.M236L, and the triple mutant p.P261L/T292A/I298T) on the biochemical properties of the polymerase in vitro, following bacterial expression and purification of the respective enzymatic variants. We report that all missense somatic pol β mutations significantly affect enzyme function. Two of the pol β variants reduce catalytic efficiency, while the remaining five missense mutations alter the fidelity of DNA synthesis. Thus, we conclude that a significant proportion (9 out of 26; 35%) of prostate cancer patients have functionally important somatic mutations of pol β . Many of these missense mutations are clonal in the tumors, and/or are associated with loss of heterozygosity and microsatellite instability. These results suggest that interfering with normal polymerase β function may be a frequent mechanism of prostate tumor progression. Furthermore, the availability of detailed structural information for pol β allows understanding of the potential mechanistic effects of these mutants on polymerase function.

Hum Mutat 32:415–423, 2011. © 2011 Wiley-Liss, Inc.

KEY WORDS: somatic; expression analysis; kinetic; POLB; mutation

Introduction

Human DNA polymerase β (pol β) is a monomeric protein of 335 residues, which is essential for short-patch base excision repair [Goodman, 2002]. Base excision repair (BER) is one of the major pathways of DNA repair that removes oxidized and alkylated bases from DNA [Friedberg, 2003]. Pol β is also involved in meiotic recombination [Kidane et al., 2010] and repair of double-stranded

DNA breaks through the process of nonhomologous end joining [Wilson and Lieber, 1999]. Targeted disruption of pol β in mice resulted in neonatal lethality (due to respiratory failure) growth retardation and apoptotic cell death in the developing nervous system [Sugo et al., 2000], suggesting a role for pol β in neurogenesis.

The small size and monomeric nature of DNA pol β make it an attractive candidate for biochemical and kinetic analysis. Moreover, the availability of a high resolution crystal structure of pol β makes it easier to identify potential functionally important residues for mechanistic studies. In addition, pol β shares many structural and mechanistic features with other DNA polymerases of known structure: for example, the mechanism of DNA polymerization follows an ordered binding of substrates, with the DNA template binding first [Beard and Wilson, 1998]. These attributes make pol β a good candidate for biochemical analysis.

Pol β expression is increased in many cancer cells [Scanlon et al., 1989] and overexpression of pol β results in aneuploidy and tumorigenesis in nude immunodeficient mice [Bergoglio et al., 2002]. Moreover, pol β heterozygous (\pm) mice exhibit increased single-stranded DNA breaks, chromosomal aberrations, and mutagenicity compared to normal animals [Cabelof et al., 2003]. Thus, both higher and lower pol β activity can result in increased β mutagenesis in vivo. Furthermore, several missense substitutions of pol β have also been shown to act as mutator mutants both in vivo and in vitro [Maitra et al., 2002; Shah et al., 2001].

Human DNA pol β is ubiquitously expressed, and the *POLB* gene (MIM# 174760) is located at chromosome band 8p11, a chromosomal region known to be lost during prostate cancer progression [Visakorpi et al., 1995]. Somatic pol β mutations were initially found in 2 out of 12 Japanese prostate cancer patients [Dobashi et al., 1994], but that analysis was done by single-strand conformation polymorphism, a technique that can miss mutations [Pearce et al., 2008].

We recently sequenced the complete coding region of the *POLB* gene for somatic mutations in 26 prostate cancer tissues. We identified 20 somatic mutations in these prostate tumors, nine of them missense [Makridakis et al., 2009]. With the exception of g.31911C>G (p.P242R), which substitutes the normal proline residue at position 242 with arginine, these substitutions were absent in lymphocyte DNA from the same patient. Many of the somatic mutations identified were prevalent in the tumors (i.e., they were present in more than half of the tumor chromosomes) [Makridakis et al., 2009], suggesting that they play an important role in tumor progression. Overall, 61% of the prostate cancer patients had somatic substitutions in pol β [Makridakis et al., 2009].

Additional Supporting Information may be found in the online version of this article.

Current address of Chang Long An is Department of Physiology & Cell Biology, University of Nevada School of Medicine, Reno, Nevada.

*Correspondence to: Nick M. Makridakis, Department of Epidemiology and Tulane Cancer Center, Tulane University, 1430 Tulane Avenue, SL-68, New Orleans, LA 70112. E-mail: nmakrida@tulane.edu

Molecular epidemiologic studies have shown that the p.P242R pol β substitution is significantly associated with decreased risk for colorectal cancer [Moreno et al., 2006]. However, the same substitution has also been associated with poorer lung cancer prognosis [Matakidou et al., 2007]. Functional biochemical studies may explain this discrepancy. In vitro, the p.P242R pol β substitution behaves as an antimutator [Hamid and Eckert, 2005], consistent with the colorectal cancer data.

Structural analyses may provide useful hints about the potential importance of some of the somatic prostate cancer mutations that we identified: nuclear magnetic resonance studies of [methyl- ^{13}C] methionine-labeled pol β have shown that methionine-236 of pol β interacts with a single-nucleotide gapped DNA template (the natural pol β substrate), and that this interaction is essential for pol β conformational activation [Bose-Basu et al., 2004]. Both methionine-236 and threonine-292 are involved in DNA template binding [Bose-Basu et al., 2004; Sawaya et al., 1997] and thus may be important for function. Both the g.25314A4T (p.M236L) and the g.32573A4G (p.T292A) substitutions (found in prostate cancer patients) [Makridakis et al., 2009] are predicted to destroy the hydrogen bond between these pol β residues and the DNA template [Sawaya et al., 1997]. Thus, both of these mutations may result in loss of activity and/or altered fidelity of DNA synthesis. Finally, proline-242 is in the “flexible loop,” a part of pol β that functions to position the primer and has been shown to contain several residues that cause a mutator phenotype when mutated [Dalal et al., 2004].

Even though we identified many somatic pol β mutations in prostate cancer, we do not know the functional effect of these mutations. In order to directly assess the potential role of the identified pol β mutations in prostate cancer progression, we biochemically characterized the effect of all the missense prostate cancer variants on both polymerase activity and fidelity of DNA replication, using mostly steady-state kinetic analysis. The data, presented here, demonstrate that these somatic pol β mutations may be important contributors to prostate cancer progression.

Furthermore, systematic biochemical characterization of the prostate cancer associated missense mutations of pol β , a monomeric polymerase highly suitable for mechanistic studies [Beard et al., 2002], provides valuable information on the molecular determinants of both polymerase activity and fidelity. In addition, the availability of detailed structural information for pol β allows structural–functional comparisons for all functionally important residues.

Materials and Methods

Nomenclature

The GenBank reference sequence accession number used for the genomic POLB sequence is AF491812.1.

Bacterial Strains and Growth Conditions

The *Escherichia coli* strain BL21 DE3 was used for pol β protein expression. *E. coli* DH5 α , BL21 (DE3), and recombinant *E. coli* harboring human pol β variants were cultured in LB medium containing 25 $\mu\text{g}/\text{ml}$ kanamycin (Sigma-Aldrich, St. Louis, MO) when appropriate.

Construction of Variants of pol β

The distinct pol β mutants were obtained by site-directed mutagenesis using the Quick-change kit (Stratagene, La Jolla, CA)

according to the protocol of the manufacturer using the pET28a(+)-WT (wild-type) bacterial expression vector as a template (a gift of Dr. Joann Sweasy from Yale University). Successful mutagenesis was confirmed by DNA sequencing with BigDye chemistry on a 3100 ABI sequencer (Perkin-Elmer, Waltham, MA).

Expression and Purification of Mutant Enzymes

Purification of pol β proteins was performed as previously described [An et al., 2004; Kosa and Sweasy, 1999] with the following modifications. Each pol β variant was expressed as a fusion protein with a six-residue poly-histidine tag at the N terminus. The enzymes were purified using HisTrap FF crude Kit (GE Healthcare, Piscataway, NJ) according to the manufacturer instructions. The fusion proteins were expressed in BL21 DE3 cells, which were grown at 37°C to mid-log phase and then induced 3 to 6 hr with 1 mM IPTG. Cells were harvested by centrifugation, resuspended in 40 mM Tris pH 8, 500 mM NaCl, 10 mM imidazole, and 100 μl Protease Inhibitor Cocktail (Sigma-Aldrich) and lysed by sonication. Extracts were cleared by centrifugation (15,000 rpm, 15 min at 4°C), and then loaded onto HisTrap FF crude Kit/100 ml of starting culture. The proteins were eluted with 500 mM imidazole in 0.5 M NaCl. Eluted proteins were then loaded onto a HiTrap SP HP column (GE Healthcare). The column was washed with 100 mM NaCl and proteins were eluted with 2 M NaCl and stored at -80°C in 50 mM Tris pH 8, 1 mM EDTA, 2 M NaCl, 10% glycerol, and protease inhibitors as above. Purified proteins were run on a Coomassie Blue-stained SDS-PAGE gel to assess purity. Protein levels were quantified by Bradford protein assay (Sigma-Aldrich).

Western Blotting

Pol β proteins were identified by Western blot [Servant et al., 2002]. Proteins were electrophoresed in a 12% SDS-PAGE gel and transfer to polyvinylidene difluoride membrane (Thermo Scientific, Waltham, MA). Blots were blocked by 5% nonfat dry milk in Tris-buffered saline–Tween 20 (0.1% Tween) and incubated with anti-His tag antibody (Sigma-Aldrich) according to the protocol of the manufacturer. For detection we were used IR Dye 800CW Goat Anti-Rabbit IgG (LI-COR Biosciences, Lincoln, NE) and the Odyssey apparatus (LI-COR Biosciences).

Assay of DNA Polymerase Activity

DNA polymerase activity assay was performed by incorporation of [α - ^{32}P]dATP (Perkin-Elmer) as previously described [Maitra et al., 2002] with the following modifications. The final reaction mixture was 50 mM Tris buffer pH 8.0, 20 mM MgCl_2 , 100 mM NaCl, 200 $\mu\text{g}/\text{ml}$ bovine serum albumin (BSA), 200 μM dithiothreitol, 20 μM dATP, 100 μM each of the three remaining dNTPs, 2 μCi of [α - ^{32}P]dATP, and 10 μg activated calf thymus DNA. Reactions were incubated at 37°C for 30 min and stopped with EDTA. The reaction mixture were spotted on GFA filters (Whatman, Piscataway, NJ), which were washed twice with 22.5 mg/ml NaPP_i , 8.5% concentrated perchloric acid and twice with 22.5 mg/ml NaPP_i , 8% concentrated hydrochloric acid, and then washed in ethanol. The filters were dried and radioactivity was counted on a scintillation counter.

DNA Substrate

All oligonucleotides were synthesized and high-pressure liquid chromatography-purified by Integrated DNA Technologies

(Coralville, IA). A 20-mer (5'-GCA GGA AAG CGA GGG TAT CC-3'), a 46-mer (5'-TAT GGT ACG CTG GAC TTT GTG GGA TAC CCT CGC TTT CCT GCT CCT G-3'), and a 20-mer (5'-ACA AAG TCC AGC GTA CCA TA-3') oligonucleotides were used as primer, template and downstream oligonucleotide, respectively [Chagovetz et al., 1997]. 5'-³²P end labeling of the 20-mer primer was performed with 3,000 Ci/mmol [γ -³²P]ATP (Perkin-Elmer) using T4 polynucleotide kinase (U.S. Biochemical Corp., Cleveland, OH) according to the manufacturer's protocol. The ³²P-labeled primer was then purified from unincorporated label by a MicrospinTM G-50 (GE Healthcare) column. The downstream oligonucleotide was 5'-end phosphorylated by Integrated DNA Technologies. The oligonucleotides were annealed at a primer: template: downstream oligonucleotide molar ratio of 1:1.2:1.3 in 50 mM Tris, pH 8.0, 250 mM NaCl, in order to create a single nucleotide gap. The mixture was incubated sequentially at 95°C for 5 min, slow cooled to 50°C for 30 min, and 50°C for 20 min and then immediately transferred to ice. Annealing of primer was confirmed on an 18% polyacrylamide (acrylamide/bis-acrylamide: 29:1) native gel followed by autoradiography as described [Li et al., 1999; Maitra et al., 2002].

Steady-State Incorporation Experiments

All incorporation reactions (20 μ l) were performed in 50 mM Tris-Cl (pH 8.0), 10 mM MgCl₂, 2 mM DTT, 20 mM NaCl, 0.2 mg/ml BSA, 2.5% glycerol and contained 50 nM annealed DNA substrate (see above). Correct incorporation (activity) reactions contained 2.5 nM purified pol β enzymes except for the triple mutant (600 nM), while misincorporation (fidelity) reactions contained 2.5 nM to 300 nM purified pol β enzymes. All concentrations given refer to the final concentrations after mixing. All reactions were performed by first preincubating the DNA substrate with pol β for 3 min at 37°C without the dNTPs. Reactions were initiated by the addition of a single dNTP (0.1–1400 μ M). After 2 min incubation at 37°C, the reactions were quenched by adding 20 μ l of formamide loading buffer (formed by mixing 900 μ l formamide, 22.2 μ l 0.5 M EDTA [pH 8.0], and 77.8 μ l water) and then boiled for 10 min, and immediately transferred into ice. Products were resolved on a 15% polyacrylamide (acrylamide/bis-acrylamide: 29:1) gel containing 7 M urea and then quantified by a Typhoon Trio+ Variable Mode PhosphorImager (GE Healthcare). To ensure that all reactions were conducted at steady state the enzyme concentrations were optimized using time course experiments [Chagovetz et al., 1997].

Kinetic Analysis

Kinetic data analyses were based on Lineweaver-Burk plots. We determined the values of k_{cat} and $K_{m,dNTP}$ from trendline equations calculated from these plots by the Microsoft Excel software (Microsoft, Redmond, WA), using values obtained from the plotted kinetic data. The $k_{cat}/K_{m,dNTP}$ values of mispairs were determined from dividing k_{cat} and K_m values obtained by Excel analysis of the inverse plots. Fidelity values for each dNTP were calculated using the following equation: Fidelity = $[(k_{cat}/K_{m,dNTP})_{correct} + (k_{cat}/K_{m,dNTP})_{incorrect}] / (k_{cat}/K_{m,dNTP})_{incorrect}$ [Li et al., 1999]. Representative plots are shown in Supp. Figure S1.

Assay of DNA pol β Lyase Activity

DNA poly β lyase activity assay was performed as in Prasad et al. [1998], with the following modifications. Briefly, the HPLC

purified DNA oligonucleotide (Integrated DNA Technologies) 49_lyase, was 3' end labeled with [α -³²P]dATP (Perkin-Elmer) by Terminal Transferase (New England Biolabs Inc., Ipswich, MA). The sequence of 49_lyase is 5'ACTACAAATTAGAAAATAGCUG-TCCTTGACGGCTAGAATTACCTACCGG3', which contains a uracil at position 21 (underlined). The 50- μ l tailing reaction includes 50 mM potassium acetate, 20 mM Tris-acetate (pH 7.9), 10 mM magnesium acetate, 0.25 mM CoCl₂, 10 μ Ci of [α -³²P]dATP and 5 pmol of substrate DNA. The reaction was incubated at 37°C for 30 min. After incubation, the labeled oligonucleotide was purified by Micro Bio-Spin chromatography Columns (Bio-Rad, Hercules, CA). The complimentary strand of 49_lyase is RC49_lyase, whose sequence is 5'CCGGTAGGTAA-TTCTAGCCGTC AAGGACAGCTATTTTCTAATTTGTAGT3', and has dATP at the uracil corresponding position. Following annealing of the forward and reverse strands, the AP site was created by removal of the uracil with Uracil DNA Glycosylase (UDG). The UDG reaction included 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM DTT, 18 nM cold annealing substrate, 3×10^4 cpm [α -³²P] dATP-labeled annealing substrate and 10 units of UDG (New England Biolabs Inc.) in a 300- μ l final volume. The UDG reaction was performed at 37°C for 30 min then stopped by phenol/chloroform extraction. Purified pol β variants were tested for their lyase activity. The reaction included 50 mM Hepes pH7.4, 2 mM DTT, 5 mM MgCl₂, with or without 0.2 units of apurinic endonuclease 1 (APE I) (New England Biolabs Inc.). Final volume was 10 μ l. The reaction was incubated at 37°C for 30 min and stopped by addition of 10 μ l 2 \times formamide loading buffer, which contains 20 mM EDTA and 95% formamide. Each assay was done in triplicate. After denaturing at 75°C for 2 min, 10 μ l sample was loaded on to 15% polyacrylamide gel containing 7 M urea. Electrophoresis was done at 75 W for 80 min in 1 \times TBE buffer by Sequi-Gen GT Electrophoresis Cell (Bio-Rad, Hercules, CA). Gel drying and autoradiography were done as usual. Bands captured and quantified with software Image Quant 5.1. The lyase activity was calculated following kinetic analysis, as described above.

Results

Expression Constructs of Variants of DNA pol β

Somatic variants of DNA pol β originally found in prostate cancer tissue [Makridakis et al., 2009] were obtained by site-directed mutagenesis of the human pol β cDNA, as described in Experimental Procedures. Both normal (WT) and variants of pol β were expressed in *E. coli* and purified using column filtration techniques, as described in Experimental Procedures. After purification, WT and variants of DNA pol β were analyzed by SDS-PAGE. This analysis showed that we obtained >90% pure versions of pol β for all variants except the triple mutant (for which we obtained 51% homogeneity; data not shown). Figure 1 displays WT pol β protein analyzed on a coomassie blue-stained SDS-polyacrylamide gel and detected by Western blot.

Variants of DNA pol β Have Polymerase Activity

We tested the polymerase activity of WT and variants of human DNA pol β by incorporating dNTPs into activated calf thymus DNA. These experiments showed that all variants of DNA pol β were active (data not shown).

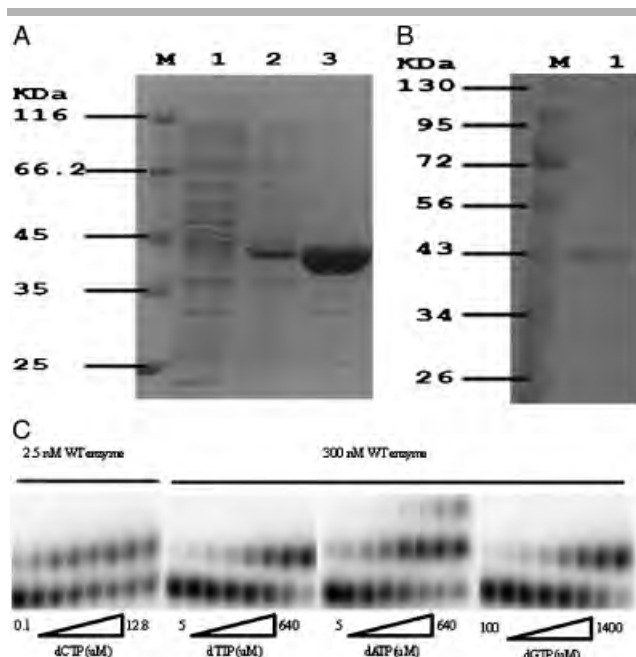


Figure 1. Identification of purified pol β protein and gel assay of pol β single nucleotide insertion. **A:** Electrophoretic analysis of the expressed pol β protein at various stages of purification. Separation was performed on a 12.5% (W/V) SDS-polyacrylamide gel. M, marker; lane 1, crude extract from BL21 (DE3) *E. coli* cells containing pET-28a(+)/pol β ; lane 2, crude extract from IPTG-induced BL21 (DE3) *E. coli* cells containing pET-28a(+)/pol β ; lane 3, purified pol β . **B:** Detection of purified protein by western blot. M, marker; lane 1, purified wild-type pol β . **C:** A gapped oligo substrate was incubated with wild-type pol β in increasing concentrations of a single dNTP as described under "Experimental Procedures." Pol β concentrations was adjusted to optimize detection of primer extension products. dATP misincorporation resulted in a single nucleotide extension of the misinserted dATP (correct extension against the next nucleotide [dT] of the DNA substrate) but only at higher dATP concentrations.

Effect of DNA pol β Variants on Catalytic Efficiency (Correct Incorporation)

To perform pol β kinetic analysis we employed a DNA substrate containing a single-nucleotide gap that was previously used for fidelity studies [Chagovetz et al., 1997; Roberts and Kunkel, 1996]. We determined pol β catalytic efficiency based on steady-state kinetic analyses of single-nucleotide addition (dCTP) opposite template dG on our single gapped DNA substrate. The catalytic activities (k_{cat}) of both the WT and genetic variants of DNA pol β are listed in Table 1. This table shows that the catalytic activity of the g.15180G4A (p.E123K) variant was increased 63% compared to WT. Variants g.25302G4A (p.E232K), p.M236L, p.P242R, g.25254G4A (p.E216K), and g.1444G4C (p.K27N) exhibit similar catalytic activity to that of WT pol β .

In comparing $K_{m,dCTP}$ between the specific variants and WT pol β (Table 1), we found that variants p.K27N and p.M236L increased the $K_{m,dCTP}$ 72% and 56%, respectively. Thus, these variants of pol β have decreased affinity for the substrate. Other variants of pol β except the triple mutant (g.32481C4T (p.P261L)/p.T292A/g.32592T4C (p.I298T); Makridakis et al. [2009]) showed similar $K_{m,dCTP}$ to that of WT pol β .

The best measure for assessing the potential effect of nucleotide variations on enzymatic activity in vivo is assaying catalytic efficiency. The catalytic efficiency ($k_{cat}/K_{m,dCTP}$) of the p.K27N variant was decreased 42% compared to WT pol β in our assays.

Table 1. Steady-State Kinetic Parameters for Correct Incorporation into a Gapped Oligo Substrate by Wild-Type pol β and Variants^a

Enzyme	k_{cat} (s^{-1}) ^b	$K_{m,dCTP}$ (μM)	$k_{cat}/K_{m,dCTP}$ ($s^{-1} M^{-1}$)
WT	$8.09 (\pm 12.06) \times 10^{-2}$	$0.25 (\pm 10.08)$	$3.24 (\pm 10.92) \times 10^5$
p.K27N	$8.06 (\pm 11.10) \times 10^{-2}$	$0.43 (\pm 10.05)$	$1.87 (\pm 10.13) \times 10^5$
p.E123K	$1.32 (\pm 10.11) \times 10^{-1}$	$0.31 (\pm 10.01)$	$4.24 (\pm 10.20) \times 10^5$
p.E232K	$1.09 (\pm 10.18) \times 10^{-1}$	$0.21 (\pm 10.02)$	$5.20 (\pm 11.35) \times 10^5$
p.P242R	$1.03 (\pm 10.21) \times 10^{-1}$	$0.27 (\pm 10.02)$	$3.81 (\pm 0.98) \times 10^5$
p.E216K	$8.62 (\pm 11.63) \times 10^{-2}$	$0.24 (\pm 10.09)$	$3.59 (\pm 1.59) \times 10^5$
p.M236L	$1.06 (\pm 10.24) \times 10^{-1}$	$0.39 (\pm 10.01)$	$2.72 (\pm 10.57) \times 10^5$
Triple mutant	$4.03 (\pm 10.27) \times 10^{-4}$	$3.72 (\pm 10.17)$	$1.08 (\pm 10.08) \times 10^2$

^aThe results represent the mean of at least three independent determinations \pm standard error. Triple mutant = p.P261L/p.T292A/p.I298T.

^bCalculated using total protein concentration.

WT, wild-type.

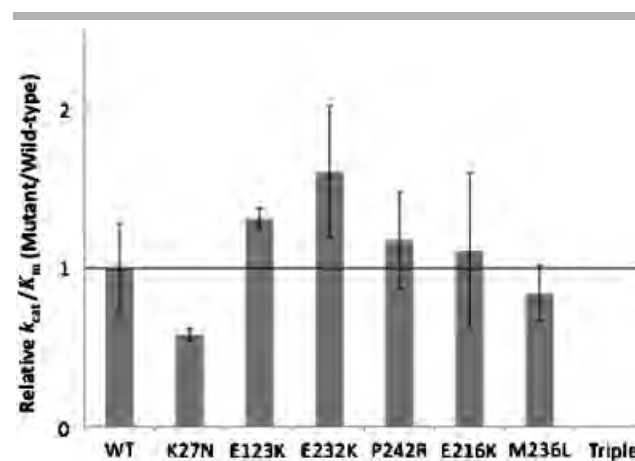


Figure 2. Influence of pol β variants on catalytic efficiency for dCTP insertion. WT and mutant variants were assayed on a single-nucleotide gapped DNA substrate with a templating dG, and the catalytic efficiency ($k_{cat}/K_{m,dNTP}$) was determined from dividing k_{cat} and K_m values obtained by Excel analysis of the inverse plots. These data represent the mean of at least three independent determinations, \pm standard error.

The remaining pol β variants, again with the exception of the triple mutant, displayed similar catalytic efficiency to that of WT pol β (Table 1 and Fig. 2).

Interestingly, the triple mutant pol β variant was significantly different than WT, or even the other variants (Table 1). The triple mutant variant showed a 99.5% decrease in catalytic activity compared with WT (Table 1). The $K_{m,dCTP}$ of the triple mutant was 15-fold higher than WT pol β and the catalytic efficiency $k_{cat}/K_{m,dCTP}$ of the triple mutant was thus dramatically decreased (Table 1 and Fig. 2) orders of magnitude lower than WT.

Effect of DNA pol β Variants on Fidelity (Misincorporation)

Misincorporation studies were performed in order to understand the role of pol β variants on DNA synthesis fidelity. These experiments were performed with enzyme concentrations that were 120-fold higher than correct incorporation experiments for both variants and WT. Figure 1C shows the results obtained for both correct and incorrect incorporation (misincorporation) experiments for the WT, with all four dNTPs.

Due to the low inherent activity of the triple pol β mutant we could not measure the catalytic activity for misincorporation even

though we used more triple mutant enzyme and dNTPs (data not shown). The k_{cat} , $K_{m,\text{dNTP}}$, and catalytic efficiency ($k_{\text{cat}}/K_{m,\text{dCTP}}$) values for misincorporation are reported in Table 2 for all variants. Percentage changes of the k_{cat} , $K_{m,\text{dNTP}}$, and catalytic efficiency ($k_{\text{cat}}/K_{m,\text{dCTP}}$) were obtained from Table 2 and listed on Table 3. Variants p.K27N and p.E123K showed similar k_{cat} to that of WT enzyme for dTTP, dATP, and dGTP misincorporations (Table 3). Variants p.P242R and p.E216K displayed decreased k_{cat} for all dNTP misincorporations (Table 3). The p.E232K variant had similar k_{cat} to WT enzyme for dATP and decreased k_{cat} for dTTP and dATP misincorporations (Table 3). The p.M236L variant showed similar k_{cat} for dTTP and dATP with the WT enzyme, and increased k_{cat} for dGTP misincorporation compared to WT (Table 3).

The K_m for dTTP misincorporation was increased for variants p.K27N, p.E216K, p.E232K, and p.P242R (Table 3). The K_m of both p.E232K and p.P242R variants was increased for both dTTP and dATP misincorporations, but the K_m for dGTP misincorporation was decreased only for the p.E232K variant, whereas the p.P242R variant showed K_m similar to WT (Table 3). The p.E123K and p.M216L variants showed WT K_m for dTTP and dGTP misincorporations and decreased K_m for dATP misincorporations (Table 3).

In compared catalytic efficiencies ($k_{\text{cat}}/K_{m,\text{dNTP}}$) for misincorporation, variants p.K27N, p.P242R, and p.E216K showed decreased efficiency for all misincorporations (Table 3). The p.E232K variant had decreased catalytic efficiency for both dTTP and dATP misincorporations but showed WT efficiency for dGTP misincorporation (Table 3). The $k_{\text{cat}}/K_{m,\text{dNTP}}$ of p.E123K and p.M236L variants was increased for dTTP and dATP misincorporations but was similar to WT enzyme for dGTP misincorporation (Table 3).

Table 2. Steady-State Kinetic Parameters for Incorrect Incorporation into a Gapped Oligo Substrate by Wild-Type pol β and Variants^a

Enzyme	k_{cat} (s^{-1}) ^b	$K_{m,\text{dCTP}}$ (μM)	$k_{\text{cat}}/K_{m,\text{dNTP}}$ ($\text{s}^{-1} \text{M}^{-1}$)
dG-dTTP			
WT	$1.45 (\pm 0.02) \times 10^{-3}$	$52.20 (\pm 1.65)$	$27.78 (\pm 0.50)$
p.K27N	$1.40 (\pm 0.03) \times 10^{-3}$	$83.78 (\pm 3.23)$	$16.71 (\pm 0.47)$
p.E123K	$1.47 (\pm 0.07) \times 10^{-3}$	$47.60 (\pm 3.73)$	$30.88 (\pm 0.99)$
p.E232K	$1.39 (\pm 0.03) \times 10^{-3}$	$59.69 (\pm 1.31)$	$23.29 (\pm 0.40)$
p.P242R	$1.39 (\pm 0.03) \times 10^{-3}$	$60.98 (\pm 1.72)$	$22.79 (\pm 0.43)$
p.E216K	$1.32 (\pm 0.03) \times 10^{-3}$	$85.21 (\pm 14.33)$	$15.49 (\pm 2.18)$
p.M236L	$1.38 (\pm 0.08) \times 10^{-3}$	$45.02 (\pm 5.79)$	$30.65 (\pm 2.30)$
dG-dATP			
WT	$1.31 (\pm 0.02) \times 10^{-3}$	$46.78 (\pm 1.90)$	$28 (\pm 1.56)$
p.K27N	$1.28 (\pm 0.04) \times 10^{-3}$	$74.59 (\pm 5.55)$	$17.16 (\pm 0.81)$
p.E123K	$1.31 (\pm 0.02) \times 10^{-3}$	$38.82 (\pm 4.65)$	$33.75 (\pm 3.86)$
p.E232K	$1.26 (\pm 0.05) \times 10^{-3}$	$457.59 (\pm 47.18)$	$2.18 (\pm 0.11)$
p.P242R	$1.22 (\pm 0.04) \times 10^{-3}$	$57.13 (\pm 2.47)$	$21.35 (\pm 1.51)$
p.E216K	$1.10 (\pm 0.09) \times 10^{-3}$	$72.44 (\pm 14.30)$	$15.18 (\pm 1.88)$
p.M236L	$1.32 (\pm 0.04) \times 10^{-3}$	$39.67 (\pm 2.44)$	$33.27 (\pm 1.08)$
dG-dGTP			
WT	$9.78 (\pm 0.20) \times 10^{-4}$	$446.41 (\pm 15.77)$	$2.19 (\pm 0.08)$
p.K27N	$8.82 (\pm 1.84) \times 10^{-4}$	$767.89 (\pm 256.68)$	$1.15 (\pm 0.13)$
p.E123K	$9.99 (\pm 0.60) \times 10^{-4}$	$457.59 (\pm 47.18)$	$2.18 (\pm 0.11)$
p.E232K	$7.35 (\pm 0.95) \times 10^{-4}$	$322.91 (\pm 104.29)$	$2.28 (\pm 0.42)$
p.P242R	$7.98 (\pm 0.15) \times 10^{-4}$	$430.51 (\pm 23)$	$1.85 (\pm 0.09)$
p.E216K	$8.29 (\pm 0.54) \times 10^{-4}$	$748.76 (\pm 88.86)$	$1.11 (\pm 0.07)$
p.M236L	$1.16 (\pm 0.04) \times 10^{-3}$	$478.25 (\pm 48.46)$	$2.43 (\pm 0.21)$

^aThe results represent the mean of at least three independent determinations \pm standard error.

^bCalculated using total protein concentration.
WT, wild-type.

Relative Fidelity of Variants

We defined fidelity as the ratio of the sum of catalytic efficiency of correct and incorrect nucleotide incorporation over the catalytic efficiency of misincorporation (Table 4). Then, relative fidelities for all variants were obtained by dividing the fidelity of each variant over WT fidelity, using the data from Table 4 (Fig. 3). The fidelity of p.K27N was similar to that of WT enzyme for all misincorporations. The fidelity of p.E123K was similar to that of WT enzyme for dATP misincorporation, but increased for dTTP and dGTP misincorporation (Table 4 and Fig. 3). The DNA synthesis fidelity of variants p.E232K, p.P242R, and p.E216K was increased for all misincorporations compared to WT (Table 4 and Fig. 3). In contrast, the p.M236L variant showed decreased fidelity for all misincorporations (Table 4 and Fig. 3).

Deoxyribose Phosphate Lyase Activity

Unlike the remaining polymerase variants analyzed here, p.K27N is part of the deoxyribose phosphate (dRP) lyase domain of pol β [Dalal et al., 2008]. Thus, we compared the dRP lyase

Table 3. Percentage Changes of k_{cat} , $K_{m,\text{dNTP}}$ and Catalytic Efficiency ($k_{\text{cat}}/K_{m,\text{dCTP}}$) Compared to Wild Type

Enzyme	k_{cat}	$K_{m,\text{dNTP}}$	$k_{\text{cat}}/K_{m,\text{dNTP}}$
dG-dTTP			
p.K27N	—	60% (\uparrow)	40% (\downarrow)
p.E123K	—	—	11% (\uparrow)
p.E232K	4% (\downarrow)	14% (\uparrow)	16% (\downarrow)
p.P242R	4% (\downarrow)	17% (\uparrow)	18% (\downarrow)
p.E216K	9% (\downarrow)	33% (\uparrow)	44% (\downarrow)
p.M236L	—	—	10% (\uparrow)
dG-dATP			
p.K27N	—	59% (\uparrow)	39% (\downarrow)
p.E123K	—	17% (\downarrow)	21% (\uparrow)
p.E232K	—	26% (\uparrow)	24% (\downarrow)
p.P242R	7% (\downarrow)	22% (\uparrow)	24% (\downarrow)
p.E216K	16% (\downarrow)	55% (\uparrow)	46% (\downarrow)
p.M236L	—	15% (\downarrow)	19% (\uparrow)
dG-dGTP			
p.K27N	—	72% (\uparrow)	47% (\downarrow)
p.E123K	—	—	—
p.E232K	25% (\downarrow)	28% (\downarrow)	—
p.P242R	18% (\downarrow)	—	16% (\downarrow)
p.E216K	15% (\downarrow)	68% (\uparrow)	49% (\downarrow)
p.M236L	19% (\uparrow)	—	—

—, not significantly different; \uparrow , significantly increased; \downarrow , significantly decreased.

Table 4. Incorrect Incorporation Fidelity of a Single dNTP Opposite G in a Gapped Oligo Substrate by Wild-Type and Variants

Enzyme	Mispair		
	G · T	G · A	G · G
Fidelity ^a			
WT	$1.16 (\pm 0.03) \times 10^4$	$1.16 (\pm 0.07) \times 10^4$	$1.48 (\pm 0.05) \times 10^5$
p.K27N	$1.12 (\pm 0.03) \times 10^4$	$1.09 (\pm 0.05) \times 10^4$	$1.63 (\pm 0.19) \times 10^5$
p.E123K	$1.37 (\pm 0.05) \times 10^4$	$1.26 (\pm 0.15) \times 10^4$	$1.95 (\pm 0.10) \times 10^5$
p.E232K	$2.23 (\pm 0.04) \times 10^4$	$2.44 (\pm 0.16) \times 10^4$	$2.29 (\pm 0.43) \times 10^5$
p.P242R	$1.67 (\pm 0.03) \times 10^4$	$1.78 (\pm 0.12) \times 10^4$	$2.06 (\pm 0.10) \times 10^5$
p.E216K	$2.32 (\pm 0.33) \times 10^4$	$2.37 (\pm 0.28) \times 10^4$	$3.23 (\pm 0.19) \times 10^5$
p.M236L	$8.91 (\pm 0.06) \times 10^3$	$8.2 (\pm 0.03) \times 10^3$	$1.12 (\pm 0.10) \times 10^5$

^aFidelity was calculated as described under "Experimental Procedures" and from Table 1 and Table 2.

WT, wild-type; —, not significantly different.

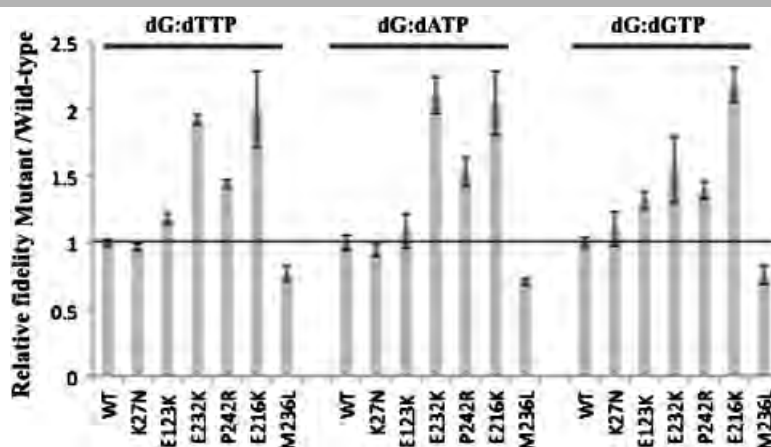


Figure 3. Relative fidelity of pol β enzyme variants. WT and mutant variants were assayed on a single-nucleotide gapped DNA substrate with a templating dG. Fidelity was calculated as described under “Experimental Procedures” and listed on Table 3. These data represent the mean of at least three independent determinations.

Table 5. Steady-State Kinetic Parameters for pol β Deoxyribose Lyase Activity

Enzyme	K_{cat} (min^{-1})	K_m (nM)	K_{cat}/K_m ($\text{nM}^{-1} \text{min}^{-1}$)
WT	4.1×10^{-3}	262	1.6×10^{-5}
p.K27N	1.1×10^{-3}	77	1.4×10^{-5}

Note: the kinetic values presented are calculated in the absence of APE I, as detailed in Materials and Methods.

activity of p.K27N to wild type, with or without prior addition of apurinic endonuclease I (APE I). The results showed that the p.K27N mutant significantly decreased both the K_m and k_{cat} , resulting in a small decrease in catalytic efficiency, K_{cat}/K_m (Table 5). A similar decrease was observed following APE I treatment (data not shown).

Discussion

We have uncovered a significant prevalence of missense nucleotide substitutions of pol β in prostate cancer tissue [Makridakis et al., 2009]. In order to assess the potential effect of these seven nucleotide substitutions (Table 1) on pol β function, we utilized expression analysis followed by kinetic assays on a DNA substrate containing a single nucleotide gap. Our results demonstrate a significant change in the catalytic activity (measured by the k_{cat}) for both the triple mutant (p.P261L/p.T292A/p.I298T) and p.E123K variants (Table 1): the triple mutant has dramatically decreased activity, while the p.E123K shows significantly increased activity. The K_m for the correct dNTP substrate (dCTP) was significantly increased (i.e., the affinity was lower) for the triple mutant, p.K27N and p.M236L variants (Table 1). The changes in pol β catalytic properties caused by these mutations resulted in significant reductions in catalytic efficiency (k_{cat}/K_m) for both the triple mutant and the p.K27N variants (Table 1 and Fig. 2). Other somatic variants (such as the p.E123K and p.E232K) showed increased catalytic efficiency compared to WT (Fig. 2), but this trend is within the experimental error of the kinetic assays.

With the exception of the triple mutant (whose expression was about half the WT level) all somatic variants resulted in essentially normal protein steady-state levels (data not shown). Thus, any

change in the observed kinetic values reflects an alteration in the catalytic properties of pol β , and not merely a change in enzyme levels. This is also true for the triple mutant, whose catalytic efficiency is reduced orders of magnitude more than the twofold decrease in enzyme levels (Fig. 2). Catalytic efficiencies at steady-state levels provide a better measure of the potential effect of the somatic mutations on prostate tumor progression rather than initial velocity or rates (k_{pol}), because they reflect the effect of these mutations over time (i.e., while the mutation is present in the tumor). As previously mentioned, both higher and lower pol β activity or levels result in increased mutagenesis in vivo. Even a 50% reduction in pol β activity or levels results in single-stranded DNA breaks, chromosomal aberrations and mutagenicity [Cabelof et al., 2003]. Based on their significant reduction on pol β catalytic efficiency (Fig. 2), both the triple mutant and the p.K27N variants are expected to result in mutagenicity, and to contribute to prostate cancer progression in vivo. The effect of the somatic mutations that marginally increase pol β activity (such as p.E123K and p.E232K; Fig. 2) is less clear. An intriguing possibility is that some of these mutations affect pol β binding to other components of the BER machinery, such as XRCC1 [Sweasy et al., 2006].

Interestingly, the p.E123K, p.E232K, and triple mutant variants are 100% prevalent in their respective prostate tumors (i.e., there is no WT allele in each patient’s tumor) [Makridakis et al., 2009]. The dramatic reduction in pol β activity conferred by the triple mutant (Fig. 2), means that the patient carrying this mutation has severely defective short-patch base excision repair in its tumor. This patient is 51 years old and has pT3a stage tumor [Makridakis et al., 2009]. Prostate cancer usually occurs in men older than 60 years [Bruner et al., 1999], suggesting that this pol β variant may contribute to early-onset prostate cancer.

The pol β knockout mouse is neonatal lethal (due to respiratory failure), shows growth retardation, and displays apoptotic cell death in the developing nervous system [Sugo et al., 2000]. The finding of a severely defective pol β variant (the triple mutant) in a prostate cancer patient’s tumor suggests that pol β activity is not essential in the adult prostate. Alternatively, it may be that this patient’s tumor has found alternative ways to compensate for the loss of pol β activity (such as activation of the long-patch base excision repair, which can also repair oxidation damage).

The triple pol β mutant (p.P261L/p.T292A/p.I298T) changes residues that are presumed to be important for both function

(T292; see Introduction) and structure: P261 forms a structurally important hydrogen bond with glutamine-264 [Pelletier et al., 1996]. P261 lies one residue away from the prostate cancer associated pol β variant p.I260M [Dalal et al., 2005]. Interestingly, the p.I260M mutant is a sequence-specific mutator that is thought to exert its effect through the disruption of the hydrogen bond between proline-261 and glutamine-264 [Dalal et al., 2005]. Thus, the p.P261L mutation of the triple mutant may also be a mutator mutant. Future biochemical analyses with the three independent mutations that compose the triple mutant may shed some light on the potential functional effects of each of these variants on base excision repair.

Furthermore, the triple pol β mutant affects an enzyme domain that has been previously shown to contain functionally important mutations: the p.E295K pol β mutation, found in gastric cancer, abolishes enzyme activity and induces cellular transformation [Lang et al., 2007]. The p.M282L pol β mutation increased mutagenesis and protein stability [Shah et al., 2001]. The p.K289M mutation, found in colon cancer, also induced mutagenesis [Lang et al., 2004]. These data suggest that the triple pol β mutant may have severe consequences in vivo.

Table 3 demonstrates that most of the somatic pol β mutations assayed here in misincorporation assays significantly change the K_m for dNTP compared to WT, but not the k_{cat} . The biggest changes on the k_{cat} for dG: dGTP are afforded by the p.E232K mutant (25% decrease) and the p.M236L mutant (19% increase) (Table 3). k_{cat} for the rest of dNTPs is affected very modestly for all mutants (less than 16%; Table 3). However, the K_m for dNTP misincorporation is increased up to 60% for dTTP, 59% for dATP and 72% for dGTP (all for the p.K27N mutant). Interestingly, most somatic mutations increase the K_m for dNTP misincorporation (Table 3). The biggest decrease in the K_m for dNTP misincorporation is attained by the p.E232K mutant: a 28% decrease in the K_m for dGTP misincorporation.

The catalytic efficiency ($k_{cat}/K_{m,dNTP}$) for dNTP misincorporation is significantly lower than WT for most of the somatic mutations (Table 3). The biggest change on the $k_{cat}/K_{m,dNTP}$ for dTTP misincorporation is achieved by the p.E216K mutant (44% decrease), whereas for dATP it is a 39% decrease (by p.K27N) and for dGTP it is a 49% decrease (by the p.E216K mutant).

Consistent with the change in catalytic efficiency for misincorporation, most somatic pol β mutations show significantly increased fidelity of DNA synthesis (Fig. 3), suggesting that they function as antimutators. The p.K27N mutation that demonstrated reduced activity for correct incorporation (Fig. 2) displays no change in fidelity compared to WT (Fig. 3). We were unable to measure fidelity for the triple mutant, due to its very low activity for misincorporation (data not shown). The p.E123K mutation displays significantly increased fidelity for dTTP and dGTP (up to 32%), but not dATP (Fig. 3). The p.E232K, p.P242R, and p.E216K mutations display significantly increased fidelity (up to 118%) for all dNTPs (Fig. 3). In contrast, the p.M236L variant showed decreased fidelity compared to WT: 23%, 29%, and 24% for dTTP, dATP, and dGTP, respectively (Table 4 and Fig. 3). Thus, the p.M236L mutation may confer a mutator phenotype. Interestingly, the prostate tumor that bares the p.M236L mutation has microsatellite instability [Makridakis et al., 2009], supporting this hypothesis. Furthermore, the effect of the p.M236L on synthesis fidelity may be due to the destruction (in the mutant) of a hydrogen bond important for template binding (see Introduction).

As mentioned above, the p.E123K and p.E232K variants are 100% prevalent in their respective prostate tumors [Makridakis et al., 2009]. These somatic variants, together with the p.E216K

(present in 52% of its respective tumor) [Makridakis et al., 2009], show significantly increased fidelity compared to WT, and thus may function as antimutators. Interestingly, all three of these variants show a trend towards higher pol β activity (although it does not reach statistical significance in our assay; Fig. 2). It is tempting to hypothesize that these three variants may actually reflect the response of prostate tumors to chemotherapy: DNA damage (e.g., alkylation) caused by chemotherapeutic drugs may actually select for tumor mutations that have both increased pol β activity and fidelity, in order to repair the damage. We do not have data on the chemotherapeutic regimen given to these patients, so we cannot directly probe this scenario at this time. However, increased pol β expression has been significantly associated with poorer chemotherapeutic response and prognosis in colorectal cancer [Iwatsuki et al., 2009]. An alternative model is that the p.E123K, p.E232K, and p.E216K pol β mutations are actually "passengers," that is, not "drivers" of tumor progression.

Several previously characterized pol β mutants exhibit misincorporation bias. For example, the p.D246V pol β mutant, present in the "flexible loop" (where the p.P242R mutant also lies; see Introduction), preferentially misincorporates dTTP opposite to templated dG [Dalal et al., 2004]. This misincorporation bias makes the p.D246V a mutator mutant mainly for C>T transitions. We examined our somatic pol β mutants for misincorporation bias. Table 4 indicates a similar trend for mutator/antimutator status for all template: dNTP misincorporations and all somatic mutations that affect fidelity of DNA replication (including the p.P242R). Thus, we conclude that these somatic mutations do not result in significant misincorporation bias. However, pol β misincorporation bias is also known to depend on sequence context. Thus, it is possible that varying the template sequence may result in distinct misincorporation bias for specific mutants. Future experiments will test this hypothesis following expression analysis of the somatic pol β mutants in vivo.

The triple pol β mutant dramatically affects the K_m for dCTP (15-fold increase). None of the residues that bind dCTP (based on the crystal structure) [Sawaya et al., 1997] is directly mutated in the triple mutant. However, D276 of pol β , binds dCTP [Sawaya et al., 1997], and is also part of an α -helix that is in close proximity with two stacked β -sheets that include two of the residues mutated in the triple mutant, p.T292 and p.I298 (Jmol; <http://molvis.sdsc.edu/fgij/fg.htm?mol=2FMS>). Mutations of these residues from threonine to alanine (at position 292) and from isoleucine to threonine (at position 298) may destabilize the local structure, perhaps reducing the stacking effect (especially the I298T mutation) and thus the interaction between the β -sheets and the α -helix containing p.D276. This, in turn, could affect the triple mutant enzyme's affinity for dCTP.

The p.K27N pol β mutant significantly decreases catalytic efficiency ($k_{cat}/K_{m,dCTP}$) without changing k_{cat} (Table 1). The p.K27N effect on the pol β catalytic efficiency can be explained by a 72% increase in the K_m for dCTP (Table 1). The p.K27N mutation is not physically close to any of the residues that bind dCTP [Sawaya et al., 1997] (Jmol; <http://molvis.sdsc.edu/fgij/fg.htm?mol=2FMS>). However, the K_m for dCTP can increase due to: (1) an increase in the dissociation constant K_d (for dCTP), (2) slower rate of dCTP insertion, or (3) decreased binding affinity for the DNA template. K27 of pol β is only 6 Angstroms away from the DNA template (Jmol; <http://molvis.sdsc.edu/fgij/fg.htm?mol=2FMS>). The p.K27N mutation abolishes a positive charge on lysine 27, which may destabilize the interaction between this residue and the negatively charged DNA template backbone, resulting in lower affinity.

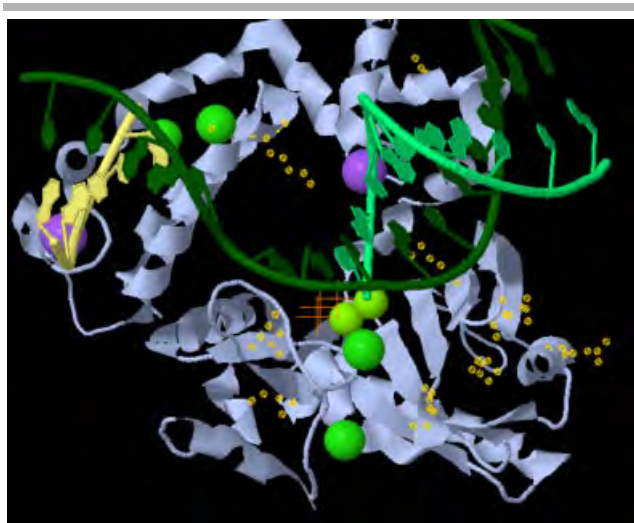


Figure 4. Location of the somatic pol β mutations. The structure of human DNA polymerase beta with dUTP (orange crosses) opposite dA and a gapped DNA substrate (2FMS) is shown using Jmol (<http://molvis.sdsc.edu/fgj/fg.htm?mol=2FMS>). Yellow “halos” indicate the side groups of the mutated residues. The gapped DNA substrate is shown in green and yellow, the polymerase β in gray. Small green spheres indicate Mg^{2+} ions, large green ones Cl^{-} ions, and purple ones Na^{+} ions.

A similar scenario may explain the effect of the triple mutant on the K_m for dCTP (alternatively explained above). The p.T292A mutation present in the triple pol β mutant is expected to abolish a hydrogen bond between this pol β residue and the DNA template (see Introduction), which may result in decreased affinity for the DNA template. This, in turn, could result in increased K_m for dCTP (as mentioned above for the K27N mutant).

Unlike the dramatic decrease on dRP lyase activity caused by the previously characterized p.L22P pol β mutant [Dalal et al., 2008], the p.K27N variant characterized here shows a small decrease in catalytic efficiency (Table 5). This finding may be due to the positioning of the side chain of K27, which points away from the lyase active site [Prasad et al., 2005].

The crystal structure of human pol β is available [Pelletier et al., 1996]. Visualization of the somatic mutations that we characterized here in pol β structure (Fig. 4) indicates that these variants are not in a specific part of the protein, but they are distributed throughout the structure. Some of the mutant residues are in areas of protein–DNA interaction, others in areas of interaction between protein domains (e.g., between two β -sheets), whereas others are in areas critical for structural maintenance [Pelletier et al., 1996]. These observations suggest that multiple structural parts of pol β are critical for BER function. The observation of several pol β mutants that affect DNA synthesis fidelity without been part of the active site is of particular interest.

In summary, we biochemically analyzed all missense somatic mutations of pol β that we previously identified in prostate cancer [Makridakis et al., 2009]. We report that all missense somatic pol β mutations have functionally significant effects: the triple mutant and the p.K27N variants affect catalytic efficiency, while the p.M236L, p.E123K, p.E232K, p.P242R, and p.E216K mutations alter the fidelity of DNA synthesis. These somatic mutations are present in a total of 7 out of 26 (27%) prostate cancer patients [Makridakis et al., 2009]. If one adds to this total the two patients with splice junction mutations (that are predicted to result in amino acid deletions) [Makridakis et al., 2009], then we conclude

that 9 of 26 (35%) of prostate cancer patients have functional somatic mutations of pol β . Functional pol β mutations have been identified at a high frequency in other types of cancer [Starcevic et al., 2004]. Thus, interfering with pol β activity may be a common mechanism of carcinogenesis. Moreover, our data significantly expands the current knowledge on the molecular determinants of both activity and fidelity for a model monomeric eukaryotic polymerase, pol β .

Acknowledgments

We thank Juergen Reichardt (Sydney) and Wanguo Liu (New Orleans) for critically reading this manuscript. N.M.M. is funded by the grant # P20RR020152-06, from the NIH.

References

- An CL, Lim WJ, Hong SY, Kim EJ, Shin EC, Kim MK, Lee JR, Park SR, Woo JG, Lim YP, Yun HD. 2004. Analysis of *bgl* operon structure and characterization of β -glucosidase from *Pectobacterium carotovorum* subsp. *carotovorum* LY34. *Biosci Biotechnol Biochem* 68:2270–2278.
- Beard WA, Shock DD, Yang XP, DeLauder SE, Wilson SH. 2002. Loss of DNA polymerase beta stacking interactions with templating purines, but not pyrimidines, alters catalytic efficiency and fidelity. *J Biol Chem* 277:8235–8242.
- Beard WA, Wilson SH. 1998. Structural insights into DNA polymerase beta fidelity: hold tight if you want it right. *Chem Biol* 5:R7–R13.
- Bergoglio V, Pillaire MJ, Lacroix-Triki M, Raynaud-Messina B, Canitrot Y, Bieth A, Garès M, Wright M, Delsol G, Loeb LA, Cazaux C, Hoffmann JS. 2002. Deregulated DNA polymerase β induces chromosome instability and tumorigenesis. *Cancer Res* 62:3511–3514.
- Bose-Basu B, DeRose EF, Kirby TW, Mueller GA, Beard WA, Wilson SH, London RE. 2004. Dynamic characterization of a DNA repair enzyme: NMR studies of [*methyl*- ^{13}C]methionine-labeled DNA polymerase β . *Biochemistry* 43: 8911–8922.
- Bruner DW, Baffoe-Bonnie A, Miller S, Diefenbach M, Tricoli JV, Daly M, Pinover W, Grumet SC, Stofey J, Ross E, Raysor S, Balslem A, Malick J, Engstrom P, Hanks GE, Mirchandani I. 1999. Prostate cancer risk assessment program. A model for the early detection of prostate cancer. *Oncology (Huntingt)* 13:325–334.
- Cabelof DC, Guo Z, Raffoul JJ, Sobol RW, Wilson SH, Richardson A, Heydari AR. 2003. Base excision repair deficiency caused by polymerase β haploinsufficiency: accelerated DNA damage and increased mutational response to carcinogens. *Cancer Res* 63:5799–5807.
- Chagovetz AM, Sweasy JB, Preston BD. 1997. Increased activity and fidelity of DNA polymerase β on single-nucleotide gapped DNA. *J Biol Chem* 272:27501–27504.
- Dalal S, Chikova A, Jaeger J, Sweasy JB. 2008. The Leu22Pro tumor-associated variant of DNA polymerase beta is dRP lyase deficient. *Nucleic Acids Res* 36:411–422.
- Dalal S, Hile S, Eckert KA, Sun KW, Starcevic D, Sweasy JB. 2005. Prostate-cancer-associated I260M variant of DNA polymerase β is a sequence-specific mutator. *Biochemistry* 44:15664–15673.
- Dalal S, Kosa JL, Sweasy JB. 2004. The D246V mutant of DNA polymerase beta misincorporates nucleotides: evidence for a role for the flexible loop in DNA positioning within the active site. *J Biol Chem* 279:577–584.
- Dobashi Y, Shuin T, Tsuruga H, Uemura H, Torigoe S, Kubota Y. 1994. DNA polymerase β gene mutation in human prostate cancer. *Cancer Res* 54: 2827–2829.
- Friedberg EC. 2003. DNA damage and repair. *Nature* 421:436–440.
- Goodman ME. 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* 71:17–50.
- Hamid S, Eckert KA. 2005. Chemotherapeutic nucleoside analogues and DNA polymerase β . *Proc Amer Assoc Cancer Res* 46:319–320.
- Iwatsuki M, Mimori K, Yokobori T, Tanaka F, Tahara K, Inoue H, Baba H, Mori M. 2009. A platinum agent resistance gene, *POLB*, is a prognostic indicator in colorectal cancer. *J Surg Oncol* 100:261–266.
- Kidane D, Jonason AS, Gorton TS, Mihaylov I, Pan J, Keeney S, de Rooij DG, Ashley T, Keh A, Liu Y, Banerjee U, Zelerman D, Sweasy JB. 2010. DNA polymerase beta is critical for mouse meiotic synapsis. *EMBO J* 29:410–423.
- Kosa JL, Sweasy JB. 1999. 3'-Azido-3'-deoxythymidine-resistant Mutants of DNA polymerase β identified by in vivo selection. *J Biol Chem* 274:3851–3858.
- Lang T, Dalal S, Chikova A, DiMaio D, Sweasy JB. 2007. The E295K DNA polymerase beta gastric cancer-associated variant interferes with base excision repair and induces cellular transformation. *Mol Cell Biol* 27:5587–5596.
- Lang T, Maitra M, Starcevic D, Li SX, Sweasy JB. 2004. A DNA polymerase beta mutant from colon cancer cells induces mutations. *Proc Natl Acad Sci USA* 101:6074–6079.

- Li SX, Vaccaro J, Sweasy JB. 1999. Involvement of phenylalanine 272 of DNA polymerase β in discriminating between correct and incorrect deoxynucleoside triphosphates. *Biochemistry* 38:4800–4808.
- Maitra M, Gudzelak Jr A, Li SX, Matsumoto Y, Eckert KA, Jager J, Sweasy JB. 2002. Threonine 79 is a hinge residue that governs the fidelity of DNA polymerase β by helping to position the DNA within the active site. *J Biol Chem* 277:35550–35560.
- Makridakis NM, Caldas Ferraz LF, Reichardt JKV. 2009. Genomic analysis of cancer tissue reveals that somatic mutations commonly occur in a specific motif. *Hum Mutat* 30:39–48.
- Matakidou A, el Galta R, Webb EL, Rudd MF, Bridle H; GELCAPS Consortium, Eisen T, Houlston RS. 2007. Genetic variation in the DNA repair genes is predictive of outcome in lung cancer. *Hum Mol Genet* 16:2333–2340.
- Moreno V, Gemignani F, Landi S, Gioia-Patricola L, Chabrier A, Blanco I, González S, Guino E, Capellà G, Canzian F. 2006. Polymorphisms in genes of nucleotide and base excision repair: risk and prognosis of colorectal cancer. *Clin Cancer Res* 12:2101–2108.
- Pearce CL, Van Den Berg DJ, Makridakis N, Reichardt JK, Ross RK, Pike MC, Kolonel LN, Henderson BE. 2008. No association between the SRD5A2 gene A49T missense variant and prostate cancer risk: lessons learned. *Hum Mol Genet* 17:2456–2461.
- Pelletier H, Sawaya MR, Wolffe W, Wilson SH, Kraut J. 1996. Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* 35:12742–12761.
- Prasad R, Batra VK, Yang XP, Krahn JM, Pedersen LC, Beard WA, Wilson SH. 2005. Structural insight into the DNA polymerase beta deoxyribose phosphate lyase mechanism. *DNA Repair (Amst)* 4:1347–1357.
- Prasad R, Beard WA, Strauss PR, Wilson SH. 1998. Human DNA polymerase beta deoxyribose phosphate lyase. Substrate specificity and catalytic mechanism. *J Biol Chem* 273:15263–15270.
- Roberts JD, Kunkel TA. 1996. The fidelity of eukaryotic DNA replication. In: DePamphilis ML, editor. *DNA replication in eukaryotic cells: concepts, enzymes and systems*. Cold Spring Harbor, NY: Cold Spring Harbor Press. p 217–247.
- Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. 1997. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 36:11205–11215.
- Scanlon KJ, Kashani-Sabet M, Miyachi H. 1989. Differential gene expression in human cancer cells resistant to cisplatin. *Cancer Invest* 7:581–587.
- Servant L, Cazaux C, Bieth A, Iwai S, Hanaoka F, Hoffmann JS. 2002. A role for DNA polymerase β in mutagenic UV lesion bypass. *J Biol Chem* 277: 50046–50053.
- Shah AM, Conn DA, Li SX, Capaldi A, Jäger J, Sweasy JB. 2001. A DNA polymerase β mutator mutant with reduced nucleotide discrimination and increased protein stability. *Biochemistry* 40:11372–11381.
- Starcevic D, Dalal S, Sweasy JB. 2004. Is there a link between DNA polymerase beta and cancer? *Cell Cycle* 3:998–1001.
- Sugo N, Aratani Y, Nagashima Y, Kubota Y, Koyama H. 2000. Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase β . *EMBO J* 19:1397–1404.
- Sweasy JB, Lang T, DiMaio D. 2006. Is base excision repair a tumor suppressor mechanism? *Cell Cycle* 5:250–259.
- Visakorpi T, Hyytinen E, Koivisto P, Tanner M, Keinänen R, Palmberg C, Palotie A, Tammela T, Isola J, Kallioniemi OP. 1995. In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* 9:401–406.
- Wilson TE, Lieber MR. 1999. Efficient processing of DNA ends during yeast nonhomologous end joining. Evidence for a DNA polymerase β (*Pol4*)-dependent pathway. *J Biol Chem* 274:23599–23609.